ESAIM: M2AN 42 (2008) 443–469
DOI: 10.1051/m2an:2008012

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

MULTIPLICATIVE SCHWARZ METHODS FOR DISCONTINUOUS GALERKIN APPROXIMATIONS OF ELLIPTIC PROBLEMS

Paola F. Antonietti[1] and Blanca Ayuso[2]

Abstract. In this paper we introduce and analyze some non-overlapping multiplicative Schwarz methods for discontinuous Galerkin (DG) approximations of elliptic problems. The construction of the Schwarz preconditioners is presented in a unified framework for a wide class of DG methods. For symmetric DG approximations we provide optimal convergence bounds for the corresponding error propagation operator, and we show that the resulting methods can be accelerated by using suitable Krylov space solvers. A discussion on the issue of preconditioning non-symmetric DG approximations of elliptic problems is also included. Extensive numerical experiments to confirm the theoretical results and to assess the robustness and the efficiency of the proposed preconditioners are provided.

Mathematics Subject Classification. 65N30, 65N55.

Received December 29, 2006. Revised August 13, 2007.
Published online April 1st, 2008.

1. Introduction

Based on a totally discontinuous finite element space, the first discontinuous Galerkin (DG) method was proposed in the seventies for the numerical approximation of hyperbolic problems by Reed and Hill [32], and independently, in the context of elliptic and parabolic equations in [3,20]. For a while, the use of DG methods was partially abandoned mainly due to the much larger number of degrees of freedom they require compared with continuous Galerkin finite element methods. However, the numerous and advantageous properties of DG methods (*e.g.*, tremendous flexibility in terms of mesh design and choice of shape functions, easy treatment of non-conforming meshes, straightforward design of *hp*-adaptivity strategies and weak approximation of boundary conditions) have motivated in recent years a renewed interest in DG approximations. In particular, the development of DG methods for elliptic problems has undergone a rapid development (see, *e.g.*, [2,4,5,7,17,18,33,36]). In spite of the active development of these methods, its practical utility is still very much limited by the size of the resulting algebraic linear systems of equations. A direct factorization of such systems might not be a viable option, and the use of iterative methods, such as Krylov space methods (*conjugate gradients* (CG), *generalized minimal residual* (GMRES), etc.) can result in a very slow convergence. As a consequence, the development of efficient solvers has started very recently to receive substantial attention, particularly, in connection with domain decomposition (DD) methods [1,12,22,27] and *multigrid* techniques [11,13,19,24,26].

Keywords and phrases. Domain decomposition methods, Schwarz preconditioners, discontinuous Galerkin methods.

[1] Dipartimento di Matematica, Università di Pavia, Via Ferrata 1, 27100 Pavia, Italy. paola.antonietti@unipv.it

[2] Istituto di Matematica Applicata e Tecnologie Informatiche - CNR, Via Ferrata 1, 27100 Pavia, Italy. blanca@imati.cnr.it

We focus here on Schwarz DD methods, that provide a very natural way to construct preconditioners that can be accelerated by Krylov space methods. The key idea is that, instead of solving one huge problem on a domain, it could be convenient (or necessary) to solve many smaller problems on single subdomains. The selection of the local problems has to be done to ensure the fast convergence of the method. To provide global interaction between the subregions and to improve the convergence in case of many subdomains, it also necessary to introduce a global coarse solver (with very few degrees of freedom per subdomain). We remark that, while Schwarz methods are by now well developed and understood for classical discretization methods [31,35,37,38] (see also [29,30] for practical applications), very few works are devoted to DG approximations [1,12,22,27], all from the last five years.

In this paper we propose and study some new multiplicative non-overlapping Schwarz preconditioners for the algebraic linear systems of equations arising from a wide class of DG approximations of elliptic problems. Following with the research started in [1], we provide a unified framework for the construction and analysis of multiplicative Schwarz preconditioners, which really share the features of the classical Schwarz methods. As in [1], we focus on the h-version of DG approximations, but here we also allow for the use of non-matching grids. In contrast to classical (conforming) discretization methods, some freedom in the definition of the local solvers arises due to the lack of continuity constraints across the element interfaces inherent to DG approximations. While in [22,27] the local solvers are defined as the restriction of the global bilinear form to each subdomain, we define the local solvers as the corresponding DG approximation of the continuous problem, but set in the subdomain. These two approaches, which coincide in the conforming case, are no longer the same in the DG framework. From our approach, the resulting local solvers turn out to be *approximate* rather than *exact*, as those used in [22,27]. This implies that a *local stability* property, that provides a one-sided measure of the approximation properties of the local bilinear forms, has to be shown.

For symmetric DG approximations, our analysis follows the abstract convergence theory of multiplicative Schwarz methods [8]. To complete the analysis, the presence of *approximate* local solvers implies the need of a technical assumption on the size of the penalty parameter which does not seem to be required in practice, as our numerical experiments indicate. For all the DG schemes stabilized by penalizing the jumps of the discrete solution across neighboring elements, we show that the energy norm of the error propagation operator is strictly less than one, and we exploit this result to prove that our preconditioner can be indeed accelerated with the GMRES iterative solver. Since the multiplicative Schwarz method is non-symmetric even for symmetric DG approximations we also present the corresponding symmetrized preconditioner and provide a bound for the condition number of the resulting preconditioned matrices, which allows us to conclude that the symmetrized multiplicative method can indeed be accelerated with the CG iterative solver. To the best of our knowledge this is the first time that a multiplicative preconditioner for DG approximations is considered. Furthermore, due to the close relation between multiplicative Schwarz and *multigrid* methods [38,41] (both are product type iterative methods), the construction and analysis presented here might throw some new light in the research of *multigrid* preconditioners for DG approximations, a field where few contributions [11,13,19,24,25] can be found so far.

Following with the research initiated in [1], we also discuss the issue of preconditioning the non-symmetric NIPG [33] and IIPG [18] approximations with multiplicative Schwarz methods. We examine the possible use of two different, but "related", existing theories for analyzing the observed convergence of the proposed product iterative methods: the abstract theory of multiplicative Schwarz methods for non-symmetric problems originally carried out by Cai and Widlund in [15], and the Eisenstat *et al.* GMRES convergence theory [21]. In both cases we provide numerical negative answers. On the one hand, while the lack of symmetry of the NIPG and IIPG schemes might in principle suggest the extension/adaptation of the abstract theory of [15], we numerically demonstrate that such a theory cannot be applied for the analysis of our preconditioners. The underlying reason is related to the "size" of the skew-symmetric part of the Schwarz operators; namely it is not a low order compact perturbation of the symmetric part as the theory of [15] requires. On the other hand, we also demonstrate that the Eisenstat *et al.* GMRES convergence theory [21], generally advocated in the analysis of Schwarz methods, cannot be applied to explain the observed convergence of the proposed preconditioners.

The construction and analysis of iterative methods for the solution of the linear systems of equations arising from non-symmetric DG methods is nowadays attracting special attention. Very recently, in [9] the authors overcome the main difficulty of this issue by considering a new *weakly over-penalized* version of the original NIPG method (see [10] for the details on the DG scheme), and prove the convergence of a *multigrid* method for this new DG discretization. By overpenalizing the method, the authors circumvent the lack of adjoint consistency, and succeed in proving the required L^2-error estimates for carrying out the analysis of the *multigrid* scheme. It is worth noticing that, for the method proposed in [10], the skew-symmetric part of the operator turns out to be a low order compact perturbation of the symmetric part. Therefore, the issue of analyzing iterative methods for the original NIPG and IIPG schemes remains an open problem. Also, the construction and analysis of Schwarz methods for this new non-symmetric DG scheme surely merit some further research.

We wish to note that, after we were submitting the paper, a new theory for non-stagnation of the GMRES method has came out [34]. The issue of exploring this less restrictive GMRES convergence theory, deserves without doubt further investigation and will be the subject of a future research.

Extensive numerical experiments are presented to confirm the convergence results, which also show the good performance and scalability (that is the independence of the convergence rate on the number of subdomains) of the proposed preconditioners. Also, we numerically compare the proposed preconditioners with the additive version studied in [1], and we show that, although the multiplicative preconditioner is less parallelizable, it is far faster than the additive one. We also address numerically the effect of the selection of the subdomain partition on the performance of our preconditioners. For that purpose we have considered two elliptic problems with a discontinuous diffusion matrix, and an anisotropic diffusion matrix. In both cases, the performance of the proposed preconditioners does not deteriorate. Furthermore, in the anisotropic case, we show that a clever ordering favours and speed up substantially the convergence of the product iterative method.

The paper is organized as follows. In Section 2 we set up our notation and briefly recall the unified framework for DG approximations of elliptic problems given in [4]. In Section 3 we provide a unified framework for the construction of the Schwarz preconditioners. The convergence analysis for symmetric DG approximations is given in Section 4; while in Section 5 we discuss the issue of preconditioning non-symmetric DG methods. Numerical experiments on conforming meshes and on meshes with hanging-nodes are provided in Section 6. Finally, in Section 7 we draw some conclusions.

2. Discontinuous Galerkin discretization

In this section, we introduce the model problem we will consider, set up some notation, and, following [4], briefly review the discontinuous finite element approximation of second order elliptic problems and the theoretical tools we shall require.

We consider the following model problem

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

where $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a smooth convex domain or a convex polygon or polyhedron and f is a given function in $L^2(\Omega)$.

Remark 2.1. We point out that, in spite of considering a simplified model problem, the results presented here also apply to more general second order elliptic operators in divergence form, with possibly discontinuous coefficients (see Sect. 6.3). Furthermore, with slightly minor changes, other kinds of boundary conditions could also be considered, *e.g.*, non-homogeneous Dirichlet and/or Neumann boundary conditions[1].

[1]In the case of Neumann boundary conditions prescribed on the whole boundary $\partial\Omega$ some compatibility conditions on the source term need to be assumed to ensure the well-posedness of the problem.

2.1. Notation

Throughout the paper, C and c are used to denote generic positive constants that may not be the same at different occurrences but that are always mesh-independent. The notation $x \approx y$ means that there exist positive constants c and C such that $c\,x \le y \le C\,x$.

Meshes and traces. Let $\mathcal{T}_h$ be a shape-regular (not necessarily matching) partition of Ω into disjoint open elements T such that $\overline{\Omega} = \cup_{T\in\mathcal{T}_h}\overline{T}$ where each $T \in \mathcal{T}_h$ is an affine image of a fixed master element $\widehat{T}$, and where $\widehat{T}$ is either the open unit d-simplex or the open unit hypercube in $\mathbb{R}^d$, $d = 2, 3$. Denoting by h_T the diameter of the element $T \in \mathcal{T}_h$, we define the mesh size $h = \max_{T\in\mathcal{T}_h}\{h_T\}$.

An *interior face* of $\mathcal{T}_h$ (if $d = 2$, "face" means "edge") is the (non-empty) interior of $\partial T^+ \cap \partial T^-$, where T^+ and T^- are two adjacent elements of $\mathcal{T}_h$, not necessarily matching. Similarly, a *boundary face* of $\mathcal{T}_h$ is the (non-empty) interior of $\partial T \cap \partial\Omega$, where T is a boundary element of $\mathcal{T}_h$. We denote by $\mathscr{E}^I$ and $\mathscr{E}^B$ the sets of all interior and boundary faces of $\mathcal{T}_h$, respectively, and set $\mathscr{E} = \mathscr{E}^I \cup \mathscr{E}^B$. We will use the convention that $\int_{\mathscr{E}} \varphi\,\mathrm{d}s = \sum_{e\in\mathscr{E}} \int_e \varphi\,\mathrm{d}s$.

We introduce the *local mesh size* function $\mathtt{h} \in L^\infty(\mathscr{E})$ defined as follows: $\mathtt{h}(\boldsymbol{x}) = \min\{h_{T^+}, h_{T^-}\}$, for $\boldsymbol{x}$ in the interior of $\partial T^+ \cap \partial T^-$, and $\mathtt{h}(\boldsymbol{x}) = h_T$, for $\boldsymbol{x}$ in the interior of $\partial T \cap \partial\Omega$. We shall refer to $\mathcal{T}_h$ as the "fine" mesh and we shall always proceed under the assumption that the local mesh size has bounded variation, *i.e.*, there exists a constant $C > 0$ such that

$$C^{-1}\, h_{T^-} \le h_{T^+} \le C h_{T^-}, \tag{2.2}$$

for all $T^\pm \in \mathcal{T}_h$ such that the interior of $T^+ \cap T^-$ in non empty. Roughly speaking, we avoid the mesh to be indefinitely refined in only one part of the domain. We also assume there exists $C > 0$ such that for all $T \in \mathcal{T}_h$ and for all $e \in \mathscr{E}$, $h_T \le C\, h_e$, where h_e is the diameter of $e \in \mathscr{E}$.

Finite element spaces. For a given partition $\mathcal{T}_h$ of Ω and an approximation order $\ell_h \ge 1$, we define the discontinuous finite element spaces V_h and $\boldsymbol{\Sigma}_h$ as

$$V_h = \{v \in L^2(\Omega) : v|_T \in \mathcal{M}^{\ell_h}(T) \;\; \forall T \in \mathcal{T}_h\}, \quad \boldsymbol{\Sigma}_h = \{\boldsymbol{\sigma} \in [L^2(\Omega)]^d : v|_T \in [\mathcal{M}^{\ell_h}(T)]^d \;\; \forall T \in \mathcal{T}_h\},$$

where $\mathcal{M}^{\ell_h}(T)$ is either the space $\mathcal{P}^{\ell_h}(T)$ of polynomials of degree at most ℓ_h on T, for T a triangle or a tetrahedron, or $\mathcal{Q}^{\ell_h}(T)$ for T a parallelogram or a parallelepiped, and where $\mathcal{Q}^{\ell_h}(T)$ is the mapping to T of $\widehat{\mathcal{Q}}^{\ell_h}(\widehat{T})$ (*i.e.*, polynomials of degree at most ℓ_h in each variable on $\widehat{T}$).

Trace operators. Let $e \in \mathscr{E}^I$ be an interior face shared by two elements T^+ and T^- with outward normal unit vectors $\boldsymbol{n}^\pm$, respectively. Denoting by $v^\pm$ and $\boldsymbol{\tau}^\pm$ the traces of piecewise smooth scalar and vector-valued functions v and $\boldsymbol{\tau}$, respectively, taken from the interior of $\partial T^\pm$, we define the following jump and *weighted* average operators

$$\begin{aligned} [\![v]\!] &= v^+\boldsymbol{n}^+ + v^-\boldsymbol{n}^-, & [\![\boldsymbol{\tau}]\!] &= \boldsymbol{\tau}^+\cdot\boldsymbol{n}^+ + \boldsymbol{\tau}^-\cdot\boldsymbol{n}^-, \\ \{\!\{v\}\!\}_\delta &= \delta v^+ + (1-\delta)v^-, & \{\!\{\boldsymbol{\tau}\}\!\}_\delta &= \delta\boldsymbol{\tau}^+ + (1-\delta)\boldsymbol{\tau}^-, \qquad \delta \in [0,1]. \end{aligned}$$

For $\delta = 1/2$, we drop the subindex and simply write $\{\!\{\cdot\}\!\}$. On a boundary face $e \in \mathscr{E}^B$, we set $\{\!\{v\}\!\}_\delta = v$, $[\![v]\!] = v\boldsymbol{n}$, $\{\!\{\boldsymbol{\tau}\}\!\}_\delta = \boldsymbol{\tau}$ and $[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}\cdot\boldsymbol{n}$.

Finally, we define the local lifting operators (see [4,14]) $r_e : [L^2(e)]^d \longrightarrow \boldsymbol{\Sigma}_h$ and $l_e : L^2(e) \longrightarrow \boldsymbol{\Sigma}_h$ by

$$\int_\Omega r_e(\boldsymbol{\phi})\cdot\boldsymbol{\tau}\,\mathrm{d}x = -\int_e \boldsymbol{\phi}\cdot\{\!\{\boldsymbol{\tau}\}\!\}\,\mathrm{d}s, \quad \int_\Omega l_e(q)\cdot\boldsymbol{\tau}\,\mathrm{d}x = -\int_e q\,[\![\boldsymbol{\tau}]\!]\,\mathrm{d}s \quad \forall\boldsymbol{\tau}\in\boldsymbol{\Sigma}_h. \tag{2.3}$$

Table 1. Numerical fluxes on interior faces, theoretical requirement on $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e$ for stability and symmetry of the corresponding bilinear form.

Method	$\widehat{u}(u)$	$\widehat{\boldsymbol{\sigma}}(\boldsymbol{\sigma}, u)$	Stability condition	Symmetry
SIPG [3]	$\{\!\{u\}\!\}$	$\{\!\{\nabla_h u\}\!\} - \alpha_e \mathtt{h}^{-1} [\![u]\!]$	$\alpha_* > \widetilde{\alpha}$	Yes
BRMPS [6]	$\{\!\{u\}\!\}$	$\{\!\{\nabla_h u\}\!\} + \alpha_e \{\!\{r_e([\![u]\!])\}\!\}$	$\alpha_* > \widetilde{\alpha}$	Yes
SIPG(δ) [36]	$\{\!\{u\}\!\}_{(1-\delta)}$	$\{\!\{\nabla_h u\}\!\}_\delta - \alpha_e \mathtt{h}^{-1} [\![u]\!]$	$\alpha_* > \widetilde{\alpha}$	Yes
NIPG [33]	$\{\!\{u\}\!\} + [\![u]\!] \cdot \boldsymbol{n}_T$	$\{\!\{\nabla_h u\}\!\} - \alpha_e \mathtt{h}^{-1} [\![u]\!]$	$\alpha_* > 0$	No
IIPG [18]	$\{\!\{u\}\!\} + 1/2\, [\![u]\!] \cdot \boldsymbol{n}_T$	$\{\!\{\nabla_h u\}\!\} - \alpha_e \mathtt{h}^{-1} [\![u]\!]$	$\alpha_* > \widetilde{\alpha}$	No
BMMPR [14]	$\{\!\{u\}\!\}$	$\{\!\{\boldsymbol{\sigma}\}\!\} + \alpha_e \{\!\{r_e([\![u]\!])\}\!\}$	$\alpha_* > \widetilde{\alpha}$	Yes
LDG [17]	$\{\!\{u\}\!\} - \boldsymbol{\beta} \cdot [\![u]\!]$	$\{\!\{\boldsymbol{\sigma}\}\!\} + \boldsymbol{\beta} \cdot [\![\boldsymbol{\sigma}]\!] - \alpha_e \mathtt{h}^{-1} [\![u]\!]$	$\alpha_* > 0$	Yes

2.2. Discontinuous Galerkin approximations

By introducing the auxiliary flux variable $\boldsymbol{\sigma} = \nabla u$, rewriting problem (2.1) as a first order system of equations, and following [4], the *primal* discontinuous Galerkin (DG) formulation reads as follows: find $u_h \in V_h$ such that,

$$\mathcal{A}(u_h, v_h) = \int_\Omega f v_h \,\mathrm{d}x \quad \forall\, v_h \in V_h, \tag{2.4}$$

where

$$\mathcal{A}(u_h, v_h) = \int_\Omega \nabla_h u_h \cdot \nabla_h v_h \,\mathrm{d}x + \int_{\mathscr{E}} [\![\widehat{u} - u_h]\!] \cdot \{\!\{\nabla_h v_h\}\!\} \,\mathrm{d}s + \int_{\mathscr{E}^I} \{\!\{\widehat{u} - u_h\}\!\} [\![\nabla_h v_h]\!] \,\mathrm{d}s \\ - \int_{\mathscr{E}} \{\!\{\widehat{\boldsymbol{\sigma}}\}\!\} \cdot [\![v_h]\!] \,\mathrm{d}s - \int_{\mathscr{E}^I} [\![\widehat{\boldsymbol{\sigma}}]\!] \{\!\{v_h\}\!\} \,\mathrm{d}s. \tag{2.5}$$

Here ∇_h denotes the elementwise application of the operator ∇, and $\widehat{u}(u_h)$ and $\widehat{\boldsymbol{\sigma}} = \widehat{\boldsymbol{\sigma}}(\boldsymbol{\sigma}_h, u_h)$ are the scalar and vector *numerical fluxes*, respectively. Their definition as suitable linear combinations of averages and jumps of u_h and $\boldsymbol{\sigma}_h$ determines the different DG methods (see [4] and Tab. 1 above). The stability of the DG methods is achieved by penalizing the jumps of u_h over each face $e \in \mathscr{E}$. As a consequence, the resulting bilinear forms contain a stabilization term $\mathcal{S}(\cdot,\cdot)$ that is either $\mathcal{S}^h(\cdot,\cdot)$ or $\mathcal{S}^r(\cdot,\cdot)$ defined as

$$\mathcal{S}^h(u_h, v_h) = \sum_{e\in\mathscr{E}} \alpha_e \int_e \mathtt{h}^{-1} [\![u_h]\!] \cdot [\![v_h]\!] \,\mathrm{d}s, \quad \mathcal{S}^r(u_h, v_h) = \sum_{e\in\mathscr{E}} \alpha_e \int_\Omega r_e([\![u_h]\!]) \cdot r_e([\![v_h]\!]) \,\mathrm{d}s \quad \forall\, u_h, v_h \in V_h, \tag{2.6}$$

where $\alpha_e > 0$ is a parameter independent of the mesh size. We define $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e$ and $\alpha^* = \max_{e\in\mathscr{E}} \alpha_e$, and assume that $\alpha^* \geq \alpha_* \geq 1$ and $\alpha^* \approx \alpha_*$.

In Table 1 we collect the definitions of the numerical fluxes on internal faces for the DG methods considered in this paper. On boundary faces, the definition modifies according to [4], Section 3.4 (in particular, for the LDG method, $\boldsymbol{\beta} = (0,0)^T$ on $\mathscr{E}^B$). To ensure stability, some of the DG methods require $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e$ sufficiently large, *i.e.*, $\alpha_* > \widetilde{\alpha}$ (see the summary in Tab. 1).

2.3. Theoretical tools

In this section we recall the basic tools we shall require in the analysis of our multiplicative Schwarz methods. From now on, since no confusion might arise, we drop the subindex h in the discrete functions belonging to V_h and $\boldsymbol{\Sigma}_h$.

We refer to [16] for a local inverse inequality that holds true for piecewise polynomials of a given order, and to [3] for a trace inequality that holds true for (regular enough) piecewise functions. We also recall that the inverse and trace inequality constants only depend on the shape regularity of the partition $\mathcal{T}_h$ and, for the inverse inequality, on the polynomial approximation degree.

For the analysis of our two-level multiplicative Schwarz methods for symmetric DG approximations we consider the norm induced by the bilinear form $\mathcal{A}(\cdot,\cdot)$, *i.e.*

$$\|v\|_{\mathcal{A}}^2 = \mathcal{A}(v,v) \quad \forall\, v \in V_h. \tag{2.7}$$

Observe that, for all the symmetric DG methods, provided the penalty parameter is taken so as to ensure the coercivity of $\mathcal{A}(\cdot,\cdot)$ (see Tab. 1), $\mathcal{A}(\cdot,\cdot)$ does indeed define an inner product and $\|\cdot\|_{\mathcal{A}}$ is a norm. Moreover, continuity and stability w.r.t. the norm $\|\cdot\|_{\mathcal{A}}$ are straightforwardly derived for all symmetric DG approximation, by using the definition of $\mathcal{A}(\cdot,\cdot)$ and the standard Cauchy-Schwarz inequality

(i) Continuity: $|\mathcal{A}(u,v)| \le \|u\|_{\mathcal{A}}\|v\|_{\mathcal{A}}$ for all $u,v \in V_h$;

(ii) Coercivity: $\mathcal{A}(v,v) = \|v\|_{\mathcal{A}}^2 > 0$ for all $v \in V_h$, $v \neq 0$.

For the non-symmetric NIPG and IIPG methods, since $\mathcal{A}(\cdot,\cdot)$ is not longer symmetric, it does not define an inner product. Therefore, we shall consider instead the inner product defined by the symmetric part of $\mathcal{A}(\cdot,\cdot)$ (and its induced norm), *i.e.*,

$$a(u,v) = \frac{\mathcal{A}(u,v)+\mathcal{A}(v,u)}{2}, \quad \|v\|_a^2 = a(v,v) \quad \forall\, u,v \in V_h. \tag{2.8}$$

We wish to stress that, although $\|v\|_a = \|v\|_{\mathcal{A}}$ for all $v \in V_h$ (since $a(v,v) = \mathcal{A}(v,v)$), when dealing with non-symmetric DG approximations, we shall rather use the notation $\|\cdot\|_a$ to emphasize that only the symmetric part $a(\cdot,\cdot)$ of $\mathcal{A}(\cdot,\cdot)$ defines an inner product. The issue of preconditioning non-symmetric DG methods is discussed in Section 5.

3. Multiplicative Schwarz methods

In this section, we present our two-level algorithms for the family of the DG methods including both symmetric and non-symmetric schemes. We start by setting some notation and introducing the assumptions on the partitions. Then we describe the two-level algorithms in an abstract general form and from the algebraic point of view.

3.1. Non-overlapping partitions, local and coarse solvers

Let $\{\mathcal{T}_{N_s}, N_s > 0\}$ be a family of partition of the domain Ω into N_s non-overlapping subdomains, *i.e.*, $\mathcal{T}_{N_s} = \{\Omega_i,\, i=1,\ldots,N_s\}$ with $\overline{\Omega} = \bigcup_{i=1}^{N_s} \overline{\Omega}_i$, and let $\{\mathcal{T}_H, H>0\}$ and $\{\mathcal{T}_h, h>0\}$ be the families of coarse and fine partitions (possibly with hanging nodes) with global mesh sizes H and h, respectively. We assume that they are nested, *i.e.*,

$$\mathcal{T}_{N_s} \subseteq \mathcal{T}_H \subseteq \mathcal{T}_h. \tag{3.1}$$

For each subdomain Ω_i of $\mathcal{T}_{N_s}$, $i=1,\ldots,N_s$, we denote by $\mathscr{E}_i$ the set of all faces of $\mathscr{E}$ (recall that $\mathscr{E}$ is the set of faces of the fine partition $\mathcal{T}_h$) belonging to $\overline{\Omega}_i$; we also set

$$\mathscr{E}_i^I = \{e \in \mathscr{E}_i : e \cap \partial\Omega_i = \emptyset\}, \qquad \mathscr{E}_i^B = \{e \in \mathscr{E}_i : e \cap \partial\Omega_i \cap \partial\Omega \neq \emptyset\}, \qquad \Gamma_i = \{e \in \mathscr{E}_i : e \subset \partial\Omega_i \setminus \partial\Omega\},$$

and observe that $\mathscr{E}_i = \mathscr{E}_i^I \cup \mathscr{E}_i^B \cup \Gamma_i$.

For $i=1,\ldots,N_s$, we define the *local spaces*

$$V_h^i = \left\{v \in L^2(\Omega_i) : v|_T \in \mathcal{M}^{\ell_h}(T) \;\; \forall\, T \in \mathcal{T}_h, T \subset \Omega_i\right\}, \quad \mathbf{\Sigma}_h^i = [V_h^i]^d,$$

where $\mathcal{M}^{\ell_h}$ is defined as before. The *prolongation* operators $R_i^T : V_h^i \longrightarrow V_h$ are defined as the classical inclusion operators from V_h^i to V_h, and the *restriction* operators R_i, are defined as the transpose of R_i^T with respect to the L^2-inner product. For vector-valued functions R_i^T and R_i are defined componentwise. Notice also that, $\boldsymbol{\Sigma}_h = R_1^T \boldsymbol{\Sigma}_h^1 \oplus \ldots \oplus R_1^T \boldsymbol{\Sigma}_h^{N_s}$ and $V_h = R_1^T V_h^1 \oplus \ldots \oplus R_{N_s}^T V_h^{N_s}$.

For $i = 1, \ldots, N_s$, we define the *local solvers* by considering the DG approximation of the model problem (2.1) but restricted to Ω_i, that is

$$-\Delta u_i = f|_{\Omega_i} \text{ on } \Omega_i, \quad u_i = 0 \text{ on } \partial\Omega_i.$$

Hence, in view of (2.5), the local bilinear forms $\mathcal{A}_i : V_h^i \times V_h^i \longrightarrow \mathbb{R}$ are defined by:

$$\mathcal{A}_i(u_i, v_i) = \int_{\Omega_i} \nabla_h u_i \cdot \nabla_h v_i \,\mathrm{d}x + \int_{\mathscr{E}_i} [\![\widehat{u}_i - u_i]\!] \cdot \{\!\{\nabla_h v_i\}\!\} \,\mathrm{d}s + \int_{\mathscr{E}_i^I} \{\!\{\widehat{u}_i - u_i\}\!\} [\![\nabla_h v_i]\!] \,\mathrm{d}s \\ - \int_{\mathscr{E}_i} \{\!\{\widehat{\boldsymbol{\sigma}}_i\}\!\} \cdot [\![v_i]\!] \,\mathrm{d}s - \int_{\mathscr{E}_i^I} [\![\widehat{\boldsymbol{\sigma}}_i]\!] \{\!\{v_i\}\!\} \,\mathrm{d}s, \quad (3.2)$$

here $\widehat{u}_i$ and $\widehat{\boldsymbol{\sigma}}_i$ are the *local* numerical fluxes. Their definition is given in terms of the corresponding definition of the global numerical fluxes $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$; those determining the global DG method. That is, on internal faces $e \in \mathscr{E}_i^I$, the definition of $\widehat{u}_i$ and $\widehat{\boldsymbol{\sigma}}_i$ coincides with that of $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$ on interior faces (see Tab. 1), and, on boundary faces $e \in \mathscr{E}_i^B \cup \Gamma_i$, they are defined as $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$ on boundary faces. Notice that, each $e \in \Gamma_i$ is in fact a boundary face for the local partition, but an interior face for the global partition. See [1] for further details on the relation between the *local* and global numerical fluxes.

From the definition of the local solvers, it turns out that our local solvers are *approximate*, in the sense that $\mathcal{A}(R_i^T u_i, R_i^T u_i) \neq \mathcal{A}_i(u_i, u_i)$. This is in contrast with the methods proposed in [12,22,27] where *exact* local solvers were employed. As a consequence, in our convergence analysis a *local stability* property (see Lem. 4.3 and Cor. 4.4 below) will be required.

The last step is the construction of the *coarse solver*. For a given approximation order ℓ_H, $0 \leq \ell_H \leq \ell_h$, the coarse spaces are defined as

$$V_h^0 \equiv V_H = \left\{ v \in L^2(\Omega) : v|_D \in \mathcal{M}^{\ell_H}(D) \quad \forall\, D \in \mathcal{T}_H \right\}, \qquad \boldsymbol{\Sigma}_h^0 \equiv \boldsymbol{\Sigma}_H = [V_h^0]^d.$$

In view of (3.1), it follows that $V_h^0 \equiv V_H \subseteq V_h$ and $\boldsymbol{\Sigma}_h^0 \equiv \boldsymbol{\Sigma}_H \subseteq \boldsymbol{\Sigma}_h$. The prolongation operator $R_0^T : V_h^0 \longrightarrow V_h$ is defined as the natural injection operator from V_h^0 to V_h, and, as before, R_0 is defined as the transpose of R_0^T with respect to the L^2-inner product. For vector-valued functions R_0^T and R_0 are defined componentwise. We define the coarse solver $\mathcal{A}_0 : V_h^0 \times V_h^0 \longrightarrow \mathbb{R}$ as the restriction of $\mathcal{A}(\cdot,\cdot)$ to $V_h^0 \times V_h^0$, *i.e.*,

$$\mathcal{A}_0(u_0, v_0) = \mathcal{A}(R_0^T u_0, R_0^T v_0) \quad \forall\, u_0, v_0 \in V_h^0. \quad (3.3)$$

We refer to [1] for further details.

3.2. Algebraic formulation and algorithmic aspects

In this section we describe our two-level multiplicative Schwarz algorithms in the variational framework given in [28] and from the algebraic point of view.

For $i = 0, \ldots, N_s$, we define the following projection-like operators

$$P_i = R_i^T \widetilde{P}_i : V_h \longrightarrow R_i^T V_h^i \subset V_h, \quad (3.4)$$

where the operators $\widetilde{P}_i : V_h \longrightarrow V_h^i$ are defined by

$$\mathcal{A}_i(\widetilde{P}_i u, v_i) = \mathcal{A}(u, R_i^T v_i) \quad \forall\, v_i \in V_h^i. \tag{3.5}$$

The coercivity of the local and coarse bilinear forms $\mathcal{A}_i(\cdot,\cdot)$ guarantees that the operators $\widetilde{P}_i$ (and therefore P_i) are well defined. The multiplicative Schwarz operator we propose is defined by

$$P_{mu} = I - (I - P_{N_s})(I - P_{N_s-1})\dots(I - P_0), \tag{3.6}$$

where $I : V_h \longrightarrow V_h$ is the identity operator. Following [8,38], the multiplicative Schwarz method consists in replacing the discrete problem $\mathcal{A}u = f$ by the equation $P_{mu}u = g$, with an appropriate right hand side g. We note that, even for the symmetric DG approximations, the corresponding operators P_{mu} are non symmetric, and therefore a suitable iterative solver as the *generalized minimal residual* (GMRES) method has to be used for solving the resulting linear systems of equations. This is in contrast with the situation encountered for the additive Schwarz operator proposed in [1] and defined as

$$P_{ad} = \sum_{i=0}^{N_s} P_i, \tag{3.7}$$

which turns out to be symmetric for all symmetric DG methods and non-symmetric for non-symmetric DG approximations. Nevertheless, following [38], a *symmetrized* version of the multiplicative Schwarz operator (3.6) can be defined as follows:

$$P_{mu}^{sym} = I - (I - P_0^*)\dots(I - P_{N_s-1}^*)(I - P_{N_s}^*)(I - P_{N_s})(I - P_{N_s-1})\dots(I - P_0), \tag{3.8}$$

where, for $i = 0,\dots,N_s$, P_i^* is the adjoint operator of P_i with respect to the inner product $\mathcal{A}(\cdot,\cdot)$ for symmetric DG methods, and with respect to $a(\cdot,\cdot)$ for the non-symmetric NIPG and IIPG approximations. It is clear that for the symmetrized multiplicative Schwarz method, a linear solver designed for symmetric linear systems as the *conjugate gradient* (CG) method can be used as an acceleration method.

Now we present our multiplicative Schwarz method from the algebraic point of view. Denoting by $\mathbf{A} \in \mathbb{R}^{n\times n}$, with $n = \dim(V_h)$, the matrix representation of the bilinear form (2.5), the algebraic formulation of (2.4) is given by $\mathbf{Au} = \mathbf{f}$, where $\mathbf{f} \in \mathbb{R}^n$ and $\mathbf{u} \in \mathbb{R}^n$ are the coefficient vectors on the right hand side of (2.4) and of the unknown u, respectively. For $i = 0,\dots,N_s$, we set $n_i = \dim(V_h^i)$. For each local space, let $\mathbf{A}_i \in \mathbb{R}^{n_i\times n_i}$ be the matrix representation of the local bilinear forms defined in (3.2) and let $\mathbf{A}_0 \in \mathbb{R}^{n_0\times n_0}$ be the matrix representation of the coarse bilinear form defined in (3.3). For $i = 0,\dots,N_s$, let $\mathbf{R}_i^T \in \mathbb{R}^{n\times n_i}$ and $\mathbf{R}_i \in \mathbb{R}^{n_i\times n}$ be the matrix representation of the prolongation and restriction operators R_i^T and R_i, respectively. Then, the matrix representation of the projection-like operators defined in (3.4) is given by

$$\mathbf{P}_i = \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i \mathbf{A} \in \mathbb{R}^{n\times n}, \quad i = 0,\dots,N_s. \tag{3.9}$$

The multiplicative Schwarz operator (3.6) can be written as the product of a suitable preconditioner, namely $\mathbf{B}_{mu}$, and $\mathbf{A}$, the former involving only the prolongation operators $\mathbf{R}_i^T$, the restriction operators $\mathbf{R}_i$ and the local operators $\mathbf{A}_i^{-1}$. More precisely,

$$\mathbf{P}_{mu} = \mathbf{I} - (\mathbf{I} - \mathbf{P}_{N_s})(\mathbf{I} - \mathbf{P}_{N_s-1})\dots(\mathbf{I} - \mathbf{P}_0) = \mathbf{B}_{mu}\mathbf{A}, \tag{3.10}$$

where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix. Analogously, for the additive Schwarz operator (3.7), we have

$$\mathbf{P}_{ad} = \sum_{i=0}^{N_s} \mathbf{P}_i = \mathbf{B}_{ad}\mathbf{A}.$$

We wish to stress that in our computations the preconditioner $\mathbf{B}_{mu}$ has not been built explicitly. In fact, the iterative solver used for the solution of the resulting linear system usually requires a routine that compute the action of the preconditioner on a generic vector. The following algorithm describes how to compute the action of the preconditioner $\mathbf{B}_{mu}$ on a generic vector $\mathbf{x} \in \mathbb{R}^n$.

Algorithm 3.1. *Action of the multiplicative Schwarz preconditioner $\boldsymbol{B}_{mu}$ on a generic vector $\boldsymbol{x} \in \mathbb{R}^n$.*

```
function z = B_mu x
    z = R_0^T A_0^{-1} R_0 x;
    for i = 1, ..., N_s
        z = z + R_i^T A_i^{-1} R_i (x - A z);
    end
```

A similar algorithm can be written for the symmetrized multiplicative Schwarz method (3.8), by taking into account that the matrix representation of the operator P_i^* is given by $\mathbf{P}_i^* = \mathbf{A}_s^{-1}\mathbf{P}_i^T\mathbf{A}_s$, where $\mathbf{A}_s$ is the matrix representation of the symmetric part of $\mathcal{A}(\cdot,\cdot)$, *i.e.*, $\mathbf{A}_s = (\mathbf{A} + \mathbf{A}^T)/2$. Clearly, whenever the underling DG scheme is symmetric, $\mathbf{A}_s = \mathbf{A}$ and the identity (3.9) gives $\mathbf{P}_i^* = \mathbf{A}^{-1}(\mathbf{R}_i^T\mathbf{A}_i^{-1}\mathbf{R}_i\mathbf{A})^T\mathbf{A} = \mathbf{P}_i$. We point out that, while the application of the additive preconditioner (3.7) does not involve any product with the original matrix $\mathbf{A}$, for the multiplicative preconditioner (3.6) (and also for its symmetrized version (3.8)) a product with $\mathbf{A}$ is required for the application of *each* local component (see Algorithm 3.1). Notice that, while the additive Schwarz preconditioner can be viewed as a block Jacobi preconditioner, the multiplicative one corresponds to a block Gauss-Seidel preconditioner. That is, the local solvers are performed sequentially, rather than in parallel, and consequently the multiplicative Schwarz preconditioner is not as parallelizable as the additive one. However, as we will show in the numerical experiments (see Sect. 6), the multiplicative Schwarz method is far faster than the additive one. Finally, we note that due to the different iterative solver required in each case (GMRES for the multiplicative preconditioner and CG for the additive preconditioner) each iteration in the multiplicative method is slightly more expensive than in the additive one.

4. Convergence analysis for symmetric DG approximations

In this section, following the abstract framework introduced in [8,38] (see also [37]), we present the analysis of the multiplicative Schwarz methods for symmetric DG approximations stabilized with the term $\mathcal{S}^h(\cdot,\cdot)$.

4.1. Preliminary results

We start by revising some key results proved in [1], for both symmetric and non-symmetric DG approximations, in the case of conforming grids. Their extension to non-matching grids is straightforward and we omit the details for the sake of brevity. We also include some preliminary results that we will need in our analysis.

Proposition 4.1 (stable decomposition [1])**.** *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a stable and consistent DG method. For any $u \in V_h$, let $u = \sum_{i=0}^{N_s} R_i^T u_i$, $u_i \in V_h^i$, $i = 0,\ldots,N_s$, where $u_0 \in V_h^0 = V_H$ is defined by*

$$u_0|_D = \frac{1}{\operatorname{meas}(D)}\int_D u\,\mathrm{d}x \qquad \forall\, D \in \mathcal{T}_H,$$

and $u_1,\ldots,u_{N_s}$ are determined (uniquely) by $u - R_0^T u_0 = R_1^T u_1 + \ldots + R_{N_s}^T u_{N_s}$. Then,

$$\sum_{i=0}^{N_s} \mathcal{A}_i(u_i,u_i) \le \alpha^* C_0^2 \mathcal{A}(u,u),$$

where $\alpha^* = \max_{e\in\mathscr{E}} \alpha_e$, *and*

$$C_0^2 = O\left(\max_{D\in\mathcal{T}_H} \frac{\mathrm{diam}(D)}{\min_{\substack{T\in\mathcal{T}_h\\T\subset D}} \mathrm{diam}(T)}\right).$$

Remark 4.2. Whenever the fine and coarse partitions are globally quasi uniform, *i.e.*, $H_D \approx H$ for all $D \in \mathcal{T}_H$, and $h_T \approx h$ for all $T \in \mathcal{T}_h$, the stable decomposition constant C_0^2 given in Proposition 4.1 can be rewritten in a more simple way as $C_0^2 = \mathcal{O}(H/h)$.

For symmetric DG approximations, Proposition 4.1 guarantees that the following (strictly positive) lower bound on the minimum eigenvalue of the additive Schwarz operator P_{ad} (3.7) can be derived

$$\mathcal{A}(P_{ad}u, u) \geq \frac{1}{\alpha^* C_0^2}\mathcal{A}(u,u) \quad \forall\, u \in V_h. \tag{4.1}$$

Next, result ensures that a local stability property which provides a one-sided measure of the approximation properties of the local bilinear forms holds true.

Lemma 4.3 (local stability [1])**.** *Let* $\mathcal{A}(\cdot,\cdot)$ *be the bilinear form of a stable and consistent DG method. Then, there exists* $C_\omega > 0$ *such that, for all* $u_i \in V_h^i$

$$\mathcal{A}(R_i^T u_i, R_i^T u_i) \leq \omega \mathcal{A}_i(u_i, u_i), \quad \omega = 1 + C_\omega, \quad i = 1, \ldots, N_s, \tag{4.2}$$

with C_ω *defined as*

$$C_\omega = C\left(C_{r_e} + \frac{1}{\sqrt{\alpha_*}}\right), \tag{4.3}$$

where C *only depends on the inverse and trace inequality constants and, for the LDG method, on an upper bound of the function* $\boldsymbol{\beta}$ *which enters into the definition of the numerical fluxes. Moreover,* $C_{r_e} \neq 0$ *only for the DG methods in which the stability term involves the local lifting operators* $r_e(\cdot)$ *defined in* (2.3).

As a consequence of Lemma 4.3, we have the following result.

Corollary 4.4. *Let* $\mathcal{A}(\cdot,\cdot)$ *be the bilinear form of a stable and consistent DG method stabilized by means of* $\mathcal{S}^h(\cdot,\cdot)$*. Then, there exists* $\overline{\alpha} > 0$ *such that if* $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e > \overline{\alpha}$*,*

$$\mathcal{A}(R_i^T u_i, R_i^T u_i) \leq \omega \mathcal{A}_i(u_i, u_i) \quad \forall\, u_i \in V_h^i, \quad i = 1, \ldots, N_s,$$

with $\omega \in (1, 2)$.

Proof. From (4.2) we get

$$\mathcal{A}(R_i^T u_i, R_i^T u_i) \leq (1 + C_\omega)\mathcal{A}_i(u_i, u_i) \quad \forall\, u_i \in V_h^i,$$

with C_ω defined as in (4.3). By hypothesis, since we are considering DG methods stabilized by means of $\mathcal{S}^h(\cdot,\cdot)$, $C_{r_e} = 0$ and $C_\omega = C/\sqrt{\alpha_*}$. Let $\overline{\alpha} \geq C^2$; then, if $\alpha_* > \overline{\alpha}$, $C_\omega < 1$, and therefore

$$\mathcal{A}(R_i^T u_i, R_i^T u_i) \leq (1 + C_\omega)\mathcal{A}_i(u_i, u_i) = \omega \mathcal{A}_i(u_i, u_i),$$

with $1 < \omega < 2$. $\square$

Remark 4.5. As in the classical Schwarz theory [8], our convergence analysis relies upon the hypothesis that the local stability constant $\omega \in (0, 2)$ (see Lem. 4.8 below). This implies that we have to restrict ourselves to all the symmetric DG approximations stabilized by means of $\mathcal{S}^h(\cdot,\cdot)$ for which, by adding a technical hypothesis on the size of the penalty parameter, we can guarantee $\omega \in (1, 2)$. We point out that such a restriction is only technical and it is not required in practice as we will show by numerical experiments in Section 6. Furthermore, for the DG methods stabilized by means of $\mathcal{S}^r(\cdot,\cdot)$, the hypothesis $\omega \in (1, 2)$ can never be satisfied even by increasing the size of the penalty parameter.

Remark 4.6. We wish to stress that, the restriction on the size parameter given in Corollary 4.4 is moderate. In fact, as stated in Lemma 4.3, the constant C can be easily seen to depend only on the trace and inverse inequality constants, and, for the LDG method, on an upper bound of the function $\boldsymbol{\beta}$ which enters into the definition of the numerical fluxes.

For symmetric DG approximations, the following *strengthened Cauchy-Schwarz inequalities* hold. For $1 \leq i,j \leq N_s$, there exist constants $0 \leq \varepsilon_{ij} \leq 1$, such that

$$|\mathcal{A}(R_i^T v_i, R_j^T v_j)| \leq \varepsilon_{ij}\mathcal{A}(R_i^T v_i, R_i^T v_i)^{1/2}\mathcal{A}(R_j^T v_j, R_j^T v_j)^{1/2} \quad \forall\, v_i \in V_h^i \;\; \forall\, v_j \in V_h^j. \tag{4.4}$$

Let $\rho(\boldsymbol{\mathcal{E}})$ be the spectral radius of $\boldsymbol{\mathcal{E}} = \{\varepsilon_{ij}\}_{1\leq i,j\leq N_s}$, it can be shown that $\rho(\boldsymbol{\mathcal{E}}) \leq 1 + N_c$, where N_c is the maximum number of adjacent subdomains that a given subdomain can have. By using the standard Cauchy-Schwarz inequality together with (4.4) and Lemma 4.3, it easily follows

$$\mathcal{A}\left(\sum_{i=1}^{N_s} P_i v, v\right) \leq \mathcal{A}\left(\sum_{i=1}^{N_s} P_i v, \sum_{i=1}^{N_s} P_i v\right)^{1/2} \mathcal{A}(v,v)^{1/2} \leq \omega\, \rho(\boldsymbol{\mathcal{E}})\mathcal{A}(v,v) \quad \forall\, v \in V_h. \tag{4.5}$$

Next, we prove a result for the projection-like operators P_i defined in (3.4). We first notice that, for all the symmetric DG approximations, the operators P_i are self-adjoint with respect to $\mathcal{A}(\cdot,\cdot)$; in fact, by using the symmetry of $\mathcal{A}(\cdot,\cdot)$ and the definition (3.4) of P_i, we get

$$\mathcal{A}(P_i u, v) = \mathcal{A}(v, R_i^T \widetilde{P}_i u) = \mathcal{A}_i(\widetilde{P}_i v, \widetilde{P}_i u) = \mathcal{A}(u, R_i^T \widetilde{P}_i v) = \mathcal{A}(u, P_i v), \qquad i = 0, \ldots, N_s. \tag{4.6}$$

Lemma 4.7. *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a symmetric DG method. For $i = 0, \ldots, N_s$, let P_i be the projection-like operators defined in (3.4). Then, for all $u, v \in V_h$,*

$$\mathcal{A}(P_i u, v) \leq \mathcal{A}(P_i u, u)^{1/2}\mathcal{A}(P_i v, v)^{1/2}, \qquad i = 0, \ldots, N_s \tag{4.7}$$

$$\mathcal{A}(P_i u, P_j v) \leq \omega\, \varepsilon_{ij}\mathcal{A}(P_i u, u)^{1/2}\mathcal{A}(P_j v, v)^{1/2}, \qquad i,j = 1, \ldots, N_s \tag{4.8}$$

where ω is the local stability constant given in Lemma 4.3 and, for $i,j = 1, \ldots, N_s$, ε_{ij} are the entries of the correlation matrix $\boldsymbol{\mathcal{E}}$ and, therefore, $\varepsilon_{ij} = 1$ if $\partial\Omega_i \cap \partial\Omega_j \neq \emptyset$ and $\varepsilon_{ij} = 0$ otherwise.

Proof. The self-adjointness (4.6) of P_i, the definitions (3.4) and (3.5) of P_i and $\widetilde{P}_i$, respectively, together with the standard Cauchy-Schwarz inequality, lead to

$$\mathcal{A}(P_i u, v) = \mathcal{A}_i(\widetilde{P}_i u, \widetilde{P}_i v) \leq \mathcal{A}_i(\widetilde{P}_i u, \widetilde{P}_i u)^{1/2}\mathcal{A}_i(\widetilde{P}_i v, \widetilde{P}_i v)^{1/2} = \mathcal{A}(u, P_i u)^{1/2}\mathcal{A}(v, P_i v)^{1/2}.$$

By using again (4.6) we reach (4.7). Similarly, the definition (3.5) of $\widetilde{P}_i$ together with the strengthened Cauchy-Schwarz inequalities (4.4), Lemma 4.3, the definition (3.5) and the identity (4.6) yield

$$\begin{aligned}\mathcal{A}(P_i u, P_j v) &= \mathcal{A}(R_i^T \widetilde{P}_i u, R_j^T \widetilde{P}_j v) \leq \varepsilon_{ij}\mathcal{A}(R_i^T \widetilde{P}_i u, R_i^T \widetilde{P}_i u)^{1/2}\mathcal{A}(R_j^T \widetilde{P}_j v, R_j^T \widetilde{P}_j v)^{1/2} \\ &\leq \omega\, \varepsilon_{ij}\mathcal{A}_i(\widetilde{P}_i u, \widetilde{P}_i u)^{1/2}\mathcal{A}_j(\widetilde{P}_j v, \widetilde{P}_j v)^{1/2} = \omega\, \varepsilon_{ij}\mathcal{A}(u, P_i u)^{1/2}\mathcal{A}(v, P_j v)^{1/2}.\end{aligned}$$

By using the identity (4.6) we reach (4.8), and the proof is complete. □

4.2. **Convergence results**

Following the abstract framework introduced in [8,38], we present our convergence results.
We start by introducing the $\mathcal{A}$-norm of the projection-like operators P_i,

$$\|P_i\|_{\mathcal{A}} = \sup_{\substack{v\in V_h\\ v\neq 0}} \frac{\|P_i v\|_{\mathcal{A}}}{\|v\|_{\mathcal{A}}}, \quad i = 0,\dots,N_s,$$

and observe that, thanks to Lemma 4.3, $\|P_i\|_{\mathcal{A}} \leq \omega$. For $i = 0,\dots,N_s$, we define, the i-th level error propagation operator by

$$E_i = (I - P_i)(I - P_{i-1})\dots(I - P_1)(I - P_0). \tag{4.9}$$

By setting $E_{-1} = I$, it is straightforward to see that, for $j = 0,\dots,N_s$,

$$E_{j-1} - E_j = P_j E_{j-1}, \tag{4.10}$$

which, after summation from $j = 0$ up to $j = i$, gives

$$I - E_i = \sum_{j=0}^{i} P_j E_{j-1} = P_0 + \sum_{j=1}^{i} P_j E_{j-1}. \tag{4.11}$$

Theorem 4.9, below, provides an estimate for $\|E_{N_s}\|_{\mathcal{A}}$ from which the convergence of the multiplicative Schwarz method (3.6) follows. For its proof we shall use the following result given in [8] which we report for the sake of completeness.

Lemma 4.8. *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a symmetric DG method stabilized by means of the operator $\mathcal{S}^h(\cdot,\cdot)$, and let $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e > \overline{\alpha}$, with $\overline{\alpha} > 0$ defined as in Corollary 4.4. Then,*

$$(2-\omega)\sum_{i=0}^{N_s} \mathcal{A}(P_i E_{i-1}v, E_{i-1}v) \leq \|v\|_{\mathcal{A}}^2 - \|E_{N_s}v\|_{\mathcal{A}}^2 \quad \forall\, v \in V_h, \tag{4.12}$$

where ω is the local stability constant given in Corollary 4.4.

Proof. The identity (4.10) and the self-adjointness of P_i give

$$\begin{aligned}
\mathcal{A}(P_i E_{i-1}v, E_{i-1}v) &= \|E_{i-1}v\|_{\mathcal{A}}^2 - \mathcal{A}(E_i v, E_{i-1}v) = \|E_{i-1}v\|_{\mathcal{A}}^2 - \mathcal{A}(E_i v, [E_i + P_i E_{i-1}]v)\\
&= \|E_{i-1}v\|_{\mathcal{A}}^2 - \|E_i v\|_{\mathcal{A}}^2 - \mathcal{A}(E_i v, P_i E_{i-1}v)\\
&= \|E_{i-1}v\|_{\mathcal{A}}^2 - \|E_i v\|_{\mathcal{A}}^2 - \mathcal{A}(E_{i-1}v, P_i E_{i-1}v) + \mathcal{A}(P_i E_{i-1}v, P_i E_{i-1}v),
\end{aligned}$$

that implies

$$2\,\mathcal{A}(P_i E_{i-1}v, E_{i-1}v) = \|E_{i-1}v\|_{\mathcal{A}}^2 - \|E_i v\|_{\mathcal{A}}^2 + \mathcal{A}(P_i E_{i-1}v, P_i E_{i-1}v).$$

The last term on the right hand side can be directly estimated by using the definition (3.4) of P_i, Corollary 4.4, the definition (3.5) of $\widetilde{P}_i$ and the self-adjointness of P_i

$$\mathcal{A}(P_i E_{i-1}v, P_i E_{i-1}v) \leq \omega\,\mathcal{A}_i(\widetilde{P}_i E_{i-1}v, \widetilde{P}_i E_{i-1}v) = \omega\,\mathcal{A}(P_i E_{i-1}v, E_{i-1}v).$$

Then, (4.12) follows after a summation from $i = 0$ up to $i = N_s$. □

We next show the main result of this section.

Theorem 4.9. *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a symmetric DG method stabilized by means of the operator $\mathcal{S}^h(\cdot,\cdot)$, and let $\alpha_* = \min_{e\in\mathscr{E}} \alpha_e > \overline{\alpha}$, with $\overline{\alpha} > 0$ defined as in Corollary 4.4. Let P_{mu} be the multiplicative Schwarz operator defined in (3.6), and let E_{N_s} be the corresponding error propagation operator. Then,*

$$\|E_{N_s}\|_{\mathcal{A}}^2 \leq 1 - \frac{2-\omega}{\alpha^* C_0^2(1+2\,\omega^2(N_c+1)^2)} < 1, \tag{4.13}$$

where C_0^2 is the stable decomposition constant given in Proposition 4.1, $\alpha^* = \max_{e\in\mathscr{E}} \alpha_e$, ω is the local stability constant given in Corollary 4.4 and N_c is the maximum number of adjacent subdomains a given subdomain might have. Hence, the spectral radius of E_{N_s} is strictly less than one, and the multiplicative Schwarz method converges.

Proof. For $i > 0$, the identity (4.11), the self-adjointness of P_i and Lemma 4.7 lead to

$$\begin{aligned}
\mathcal{A}(P_i v, v) &= \mathcal{A}(P_i v, E_{i-1}v) + \mathcal{A}(P_i v, P_0 v) + \sum_{j=1}^{i-1} \mathcal{A}(P_i v, P_j E_{j-1} v)\\
&= \mathcal{A}(v, P_i E_{i-1}v) + \mathcal{A}(v, P_i P_0 v) + \sum_{j=1}^{i-1} \mathcal{A}(P_i v, P_j E_{j-1} v)\\
&\leq \mathcal{A}(P_i v, E_{i-1}v) + \mathcal{A}(P_i v, P_0 v) + \omega \sum_{j=1}^{i-1} \varepsilon_{ij} \mathcal{A}(P_i v, v)^{1/2} \mathcal{A}(P_j E_{j-1} v, E_{j-1}v)^{1/2}\\
&\leq \mathcal{A}(P_i v, v)^{1/2} \left[\mathcal{A}(P_i E_{i-1} v, E_{i-1}v)^{1/2} + \mathcal{A}(P_i P_0 v, P_0 v)^{1/2} + \omega \sum_{j=1}^{i-1} \varepsilon_{ij} \mathcal{A}(P_j E_{j-1} v, E_{j-1}v)^{1/2} \right]\\
&\leq \mathcal{A}(P_i v, v)^{1/2} \left[\mathcal{A}(P_i P_0 v, P_0 v)^{1/2} + \omega \sum_{j=1}^{i} \varepsilon_{ij} \mathcal{A}(P_j E_{j-1} v, E_{j-1}v)^{1/2} \right],
\end{aligned}$$

where, in the last step, we have used that $\omega \geq 1$. Now, cancel the common factor, square both sides and use the inequality $(a+b)^2 \leq 2\,a^2 + 2\,b^2$ to obtain

$$\mathcal{A}(P_i v, v) \leq 2\,\mathcal{A}(P_i P_0 v, P_0 v) + 2\,\omega^2 \left[\sum_{j=1}^{i} \varepsilon_{ij} \mathcal{A}(P_j E_{j-1} v, E_{j-1}v)^{1/2} \right]^2.$$

Next, by summing over $i = 1, \ldots, N_s$ and extending the sum on j up to N_s (note that it will only increase the right hand side), we get

$$\sum_{i=1}^{N_s} \mathcal{A}(P_i v, v) \leq 2 \sum_{i=1}^{N_s} \mathcal{A}(P_i P_0 v, P_0 v) + 2\,\omega^2 \sum_{i=1}^{N_s} \left[\sum_{j=1}^{N_s} \varepsilon_{ij} \mathcal{A}(P_j E_{j-1} v, E_{j-1}v)^{1/2} \right]^2. \tag{4.14}$$

The first term on the right hand side can be bounded directly by using estimate (4.5):

$$2 \sum_{i=1}^{N_s} \mathcal{A}(P_i P_0 v, P_0 v) \leq 2\,\omega^2 \rho(\boldsymbol{\mathcal{E}}) \mathcal{A}(P_0 v, P_0 v) \leq 2\,\omega^2 \rho(\boldsymbol{\mathcal{E}})^2 \mathcal{A}(P_0 v, P_0 v),$$

where we have also used that $\rho(\boldsymbol{\mathcal{E}}) \geq 1$, so that $\rho(\boldsymbol{\mathcal{E}})^2 \geq \rho(\boldsymbol{\mathcal{E}})$. Denoting by $\boldsymbol{x}$ the vector of components $\mathcal{A}(P_j E_{j-1} v, E_{j-1} v)^{1/2}$, the second term on the right hand side in (4.14) can be written as $2\,\omega^2 (\boldsymbol{\mathcal{E}}\boldsymbol{x})^T (\boldsymbol{\mathcal{E}}\boldsymbol{x})$, and therefore, it is straightforward estimated by noting that $(\boldsymbol{\mathcal{E}}\boldsymbol{x})^T(\boldsymbol{\mathcal{E}}\boldsymbol{x}) \leq \rho(\boldsymbol{\mathcal{E}})^2\boldsymbol{x}^T\boldsymbol{x}$. Finally, by adding the term $\mathcal{A}(P_0v,v) = \mathcal{A}(P_0E_{-1}v, E_{-1}v)$ to both sides of (4.14), we obtain

$$\sum_{i=0}^{N_s} \mathcal{A}(P_iv,v) \leq (1+2\,\omega^2\rho(\boldsymbol{\mathcal{E}})^2)\sum_{i=0}^{N_s}\mathcal{A}(P_jE_{j-1}v, E_{j-1}v).$$

Estimate (4.1) together with the upper bound (4.12) from Lemma 4.8 give

$$\frac{1}{\alpha^* C_0^2}\,\|v\|_{\mathcal{A}}^2 \leq \sum_{i=0}^{N_s}\mathcal{A}(P_iv,v) \leq \frac{1+2\,\omega^2\rho(\boldsymbol{\mathcal{E}})^2}{2-\omega}\left(\|v\|_{\mathcal{A}}^2 - \|E_{N_s}v\|_{\mathcal{A}}^2\right).$$

Then, (4.13) follows from the estimate $\rho(\boldsymbol{\mathcal{E}}) \leq N_c + 1$. □

Remark 4.10. Theorem 4.9 guarantees that the multiplicative Schwarz preconditioner can indeed be accelerated with the GMRES method. Indeed, it is easy to see that the two sufficient conditions of the GMRES convergence theory of Eisenstat *et al.* (see [21]), usually advocated in the analysis of Schwarz methods, are straightforwardly satisfied. According to [21], the GMRES method applied to the preconditioned system does not stagnate (*i.e.*, the iterative method makes some progress in reducing the residual at each iteration step) provided that the symmetric part of P_{mu} is positive definite and P_{mu} has a bounded norm. Hence, by setting

$$c_p(P_{mu}) = \inf_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}(v, P_{mu}v)}{\mathcal{A}(v,v)}, \quad C_p(P_{mu}) = \sup_{\substack{v\in V_h\\ v\neq 0}} \frac{\|P_{mu}v\|_{\mathcal{A}}}{\|v\|_{\mathcal{A}}}, \tag{4.15}$$

we have to show that $c_p(P_{mu}) > 0$ and $C_p(P_{mu})$ is bounded. In view of Theorem 4.9, the upper bound for $C_p(P_{mu})$ follows directly from its definition:

$$\|P_{mu}\|_{\mathcal{A}} = \|I - E_{mu}\|_{\mathcal{A}} \leq 1 + \|E_{mu}\|_{\mathcal{A}}.$$

To prove the positive definiteness of the operator P_{mu}, we have to show that there exists a constant $K_0 > 0$ such that

$$\mathcal{A}(P_{mu}v,v) = \mathcal{A}((I-E_{N_s})v,v) \geq K_0\,\mathcal{A}(v,v) \quad \forall\, v\in V_h. \tag{4.16}$$

By the definition of $\|E_{N_s}\|_{\mathcal{A}}$, we have $\mathcal{A}(E_{N_s}v,v) \leq \|E_{N_s}\|_{\mathcal{A}}\,\mathcal{A}(v,v)$, and so

$$\mathcal{A}(v,v) - \mathcal{A}(E_{N_s}v,v) \geq (1-\|E_{N_s}\|_{\mathcal{A}})\;\mathcal{A}(v,v).$$

Then, (4.16) holds true with $K_0 = 1 - \|E_{N_s}\|_a$ which is positive thanks to Theorem 4.9.

As a direct consequence of Theorem 4.9, following [38], we can guarantee the convergence of the symmetrized multiplicative Schwarz method.

Corollary 4.11. *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a symmetric DG method stabilized by means of the operator $\mathcal{S}^h(\cdot,\cdot)$, and let $\alpha_* = \min_{e\in\mathscr{E}}\alpha_e > \overline{\alpha}$, with $\overline{\alpha} > 0$ defined as in Corollary 4.4. Let P_{mu}^{sym} be the symmetrized multiplicative Schwarz method defined in (3.8), and let $E_{N_s}^{sym}$ be the corresponding error propagation operator. Then,*

$$\|E_{N_s}^{sym}\|_{\mathcal{A}} \leq 1 - \frac{2-\omega}{\alpha^* C_0^2(1+2\,\omega^2(N_c+1)^2)} < 1,$$

where C_0^2 is the stable decomposition constant given in Proposition 4.1, $\alpha^* = \max_{e\in\mathscr{E}}\alpha_e$, ω is the local stability constant given in Corollary 4.4 and N_c denotes the maximum number of adjacent subdomains that a given subdomain might have.

Proof. Let $E^*_{N_s}$ bet the adjoint of E_{N_s} with respect to the $\mathcal{A}(\cdot,\cdot)$-inner product, and let $E^{sym}_{N_s}$ be the error propagation operator corresponding to P^{sym}_{mu}. Then, the proof easily follows by observing that $E^{sym}_{N_s} = E^*_{N_s} E_{N_s}$ is symmetric with respect to the $\mathcal{A}(\cdot,\cdot)$ inner product, and so

$$\|E^{sym}_{N_s}\|_{\mathcal{A}} = \|E^*_{N_s} E_{N_s}\|_{\mathcal{A}} = \|E_{N_s}\|^2_{\mathcal{A}}. \qquad \square$$

The last result guarantees that there is no qualitative difference in the convergence properties between the multiplicative Schwarz method and its symmetrized version.

Remark 4.12. Since P^{sym}_{mu} is self-adjoint with respect to $\mathcal{A}(\cdot,\cdot)$, we can use the Rayleigh quotient characterization of the extreme eigenvalues, *i.e.*,

$$\lambda^{\mathcal{A}}_{\min}(P^{sym}_{mu}) = \min_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}(P^{sym}_{mu}v,v)}{\mathcal{A}(v,v)}, \quad \lambda^{\mathcal{A}}_{\max}(P^{sym}_{mu}) = \max_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}(P^{sym}_{mu}v,v)}{\mathcal{A}(v,v)}.$$

The spectral condition number of P^{sym}_{mu} is given by $\kappa(P^{sym}_{mu}) = \lambda^{\mathcal{A}}_{\max}(P^{sym}_{mu})/\lambda^{\mathcal{A}}_{\min}(P^{sym}_{mu})$. Corollary 4.11 allows us to provide the following bound on the condition number for the symmetrized multiplicative Schwarz method:

$$\kappa(P^{sym}_{mu}) \leq \frac{\alpha^* C_0^2\left(1+2\,\omega^2(N_c+1)^2\right)}{2-\omega}. \tag{4.17}$$

In fact, an upper bound for $\lambda^{\mathcal{A}}_{\max}(P^{sym}_{mu})$ follows from the definition of P^{sym}_{mu} and the fact that $\mathcal{A}(\cdot,\cdot)$ is positive definite

$$\lambda^{\mathcal{A}}_{\max}(P^{sym}_{mu}) = \max_{\substack{u\in V_h\\ u\neq 0}} \frac{\mathcal{A}((I-E^*_{N_s}E_{N_s})v,v)}{\mathcal{A}(v,v)} = 1 - \min_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}(E_{N_s}v,E_{N_s}v)}{\mathcal{A}(v,v)} \leq 1.$$

A (strictly positive) lower bound for $\lambda^{\mathcal{A}}_{\min}(P^{sym}_{mu})$ can be proved by using Corollary 4.11

$$\lambda^{\mathcal{A}}_{\min}(P^{sym}_{mu}) = \min_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}((I-E^*_{N_s}E_{N_s})v,v)}{\mathcal{A}(v,v)} = 1 - \max_{\substack{v\in V_h\\ v\neq 0}} \frac{\mathcal{A}(E_{N_s}v,E_{N_s}v)}{\mathcal{A}(v,v)} \geq \frac{2-\omega}{\alpha^* C_0^2(1+2\,\omega^2(N_c+1)^2)}.$$

The above estimate combined with the previous one gives (4.17).

5. The issue of preconditioning non-symmetric DG methods

In this section we discuss the issue of preconditioning non-symmetric DG approximations of the model problem (2.1) with the multiplicative Schwarz preconditioner introduced in Section 3. Since our discrete bilinear forms are no longer symmetric, one might consider the extension or "adaptation" of the available theory of multiplicative Schwarz methods for non-symmetric problems originally carried out by Cai and Widlund in [15]. We shall demonstrate by means of numerical computations not only that such a theory cannot be applied, but also that the GMRES convergence theory of Eisenstat *et al.* [21] fails to explain the convergence of the proposed preconditioners. The results we present are inspired by and complete those in [1].

We start by recalling the primal bilinear form of the non-symmetric NIPG and IIPG methods:

$$\mathcal{A}(u,v) = \int_\Omega \nabla_h u\cdot\nabla_h v\,\mathrm{d}x - (1-\gamma)\int_{\mathscr{E}} [\![u]\!]\cdot\{\!\{\nabla_h v\}\!\}\,\mathrm{d}s - \int_{\mathscr{E}} \{\!\{\nabla_h u\}\!\}\cdot[\![v]\!]\,\mathrm{d}s + \mathcal{S}^h(u,v) \quad \forall\, u,v\in V_h,$$

where $\gamma = 2, 1$ for the NIPG and IIPG methods, respectively, and where $\mathcal{S}^h(\cdot,\cdot)$ is defined as in (2.6). The corresponding symmetric and skew-symmetric parts are given by

$$\begin{aligned} a(u,v) &= \int_\Omega \nabla_h u \cdot \nabla_h v \, \mathrm{d}x + \frac{(\gamma-2)}{2}\left\{\int_{\mathscr{E}} [\![u]\!] \cdot \{\!\{\nabla_h v\}\!\} \, \mathrm{d}s + \int_{\mathscr{E}} \{\!\{\nabla_h u\}\!\} \cdot [\![v]\!] \, \mathrm{d}s\right\} + \mathcal{S}^h(u,v) \quad \forall\, u,v \in V_h, \\ s(u,v) &= \frac{\gamma}{2}\int_{\mathscr{E}} [\![u]\!] \cdot \{\!\{\nabla_h v\}\!\} \, \mathrm{d}s - \frac{\gamma}{2}\int_{\mathscr{E}} \{\!\{\nabla_h u\}\!\} \cdot [\![v]\!] \, \mathrm{d}s \qquad \forall\, u,v \in V_h, \end{aligned}$$

respectively. We consider the inner product defined by the symmetric part $a(\cdot,\cdot)$ and its induced norm $\|\cdot\|_a^2 = a(\cdot,\cdot)$ (see (2.8)). The following bound can be easily shown for the skew-symmetric part (see [1]),

$$|s(u,v)| \le \frac{\gamma}{2}\left|\int_{\mathscr{E}} [\![u]\!] \cdot \{\!\{\nabla_h v\}\!\} \, \mathrm{d}s\right| + \frac{\gamma}{2}\left|\int_{\mathscr{E}} \{\!\{\nabla_h u\}\!\} \cdot [\![v]\!] \, \mathrm{d}s\right| \le C_{ss}\frac{\gamma}{\sqrt{\alpha_*}}\|u\|_a \|v\|_a, \tag{5.1}$$

where C_{ss} is a constant depending only on the shape regularity of the mesh and the polynomial degree.

Remark 5.1. We have already noticed that $\|v\|_a = \|v\|_{\mathcal{A}}$ for all $v \in V_h$. However, to emphasize that only the symmetric part of $\mathcal{A}(\cdot,\cdot)$ defines an inner product, throughout this section we will denote the norm by $\|\cdot\|_a$.

5.1. **Applicability of the abstract theory of Cai and Widlund [15]**

The general framework developed in [15] provides an upper bound (strictly less than one) of the norm of the operator E_{mu} by studying the positive definiteness of the symmetrized multiplicative operator P_{mu}^{sym}. Such a bound implies that the spectral radius of the error propagation operator, $\rho(E_{N_s})$, is strictly less than one, and therefore, a simple Richardson iteration applied to the preconditioned system converges. More precisely, by recalling that P_i^* is the adjoint operator of P_i with respect to the inner product defined by $a(\cdot,\cdot)$ (*i.e.*, $a(P_i u, v) = a(u, P_i^* v)\ \forall\, u,v \in V_h$), the symmetrized operator $\overline{P}_i$ is generally defined as

$$\overline{P}_i = P_i + P_i^* - P_i^* P_i = (I - P_i^*)(I - P_i), \quad i = 0,\dots,N_s.$$

From the definition (4.9) of the i-th level error propagation operator, it is straightforward to see that

$$I - E_{N_s}^* E_{N_s} = \sum_{i=0}^{N_s} E_{i-1}^* \overline{P}_i E_{i-1} = P_{mu}^{sym}. \tag{5.2}$$

Then, thanks to the above identity, to prove that $\|E_{N_s}\|_a < 1$, it is enough to show that the symmetrized multiplicative operator is "sufficiently" positive definite. In view of the results shown in [1], we have done some numerical computations demonstrating that the minimum eigenvalue of P_{mu}^{sym}, *i.e.*,

$$\lambda_{\min}^a(P_{mu}^{sym}) = \min_{\substack{v\in V_h \\ v\neq 0}} \frac{a(P_{mu}^{sym}v, v)}{a(v,v)},$$

might be negative in non pathological cases. In Table 2 we show the estimates of $\lambda_{\min}^a(P_{mu}^{sym})$ obtained on unstructured triangular meshes with $\ell_h = \ell_H = 1$, and with $\alpha_e = \alpha = 1$ (left) and $\alpha_e = \alpha = 10$ (right) for all $e \in \mathscr{E}$. More computations (not reported here, for brevity) that confirm the results reported in Table 2, were done for different mesh configurations, different polynomial approximation degrees and higher values of the penalty parameter. From identity (5.2), whenever $\lambda_{\min}^a(P_{mu}^{sym}) < 0$, it follows that $\|E_{N_s}\|_a > 1$, which would imply that we cannot conclude that $\rho(E_{N_s}) < 1$.

A closer inspection reveals that the theory is far from being optimal mainly due to two reasons. The first one is related to the bounds for the convergence rates of GMRES which are far from being sharp. The second relies on a required assumption on the "size" of the non-symmetric part of the problem, namely $s(\cdot,\cdot)$. In particular,

TABLE 2. Estimate of $\lambda_{\min}^a(P_{mu}^{sym})$. NIPG method: $\ell_h = \ell_H = 1$, $N_s = 16$, unstructured triangular grids.

	$\alpha = 1$			$\alpha = 10$		
$H \downarrow \ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	h_0	$h_0/2$	$h_0/4$
H_0	–1.2651	–2.6841	–3.8467	0.4357	0.0696	–0.1030
$H_0/2$	–	–1.4082	–3.5500	–	0.3937	0.1100
$H_0/4$	–	–	–1.5699	–	–	0.3966

TABLE 3. Estimate of $c_p(P_{mu})$. NIPG method: $\ell_h = \ell_H = 1$, unstructured triangular grids.

	$\alpha = 1$				$\alpha = 10$			
$H \downarrow \ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	0.1583	–0.0948	–0.2744	–0.3797	0.3370	0.1124	0.0190	–0.0237
$H_0/2$	–	0.0970	–0.2104	–0.3828	–	0.3427	0.1273	0.0383
$H_0/4$	–	–	0.0883	–0.1956	–	–	0.3278	0.1208
$H_0/8$	–	–	–	0.0254	–	–	–	0.2927

to apply the theory it is necessary to show that the non-symmetric part of the bilinear form $s(\cdot,\cdot)$ is "small" with respect to the symmetric, positive definite part $a(\cdot,\cdot)$. In our case, from estimate (5.1), one might deduce that the skew-symmetric and symmetric parts are about the same order, unless some extra conditions were imposed on the penalty parameter (*e.g.*, by taking the penalty parameter as a function of h^{-1}).

Remark 5.2. Similarly, one can conclude that the approach introduced in [39,40] does not apply to this problem.

5.2. Applicability of the Eisenstat *et al.* GMRES convergence theory [21]

Since in the numerical experiments our multiplicative Schwarz method is accelerated with the GMRES iterative solver, to develop the convergence analysis of the proposed preconditioners one might consider instead the direct application of one of the available GMRES convergence bounds. The estimate by Eisenstat *et al.* [21] is particularly well suited for Schwarz methods [37]. Such a bound guarantees the non-stagnation of the GMRES method provided the symmetric part of the operator is positive definite and the norm of the operator is uniformly bounded. More precisely, by setting as in (4.15)

$$c_p(P_{mu}) = \inf_{\substack{v \in V_h \\ v \neq 0}} \frac{a(v, P_{mu}v)}{a(v,v)}, \qquad\qquad C_p(P_{mu}) = \sup_{\substack{v \in V_h \\ v \neq 0}} \frac{\|P_{mu}v\|_a}{\|v\|_a}.$$

One needs to ensure that $c_p(P_{mu}) > 0$ and that $C_p(P_{mu})$ is bounded. While the second condition can be easily proved, the first one cannot be guaranteed. Indeed, as we demonstrate by numerical computations, $c_p(P_{mu})$ might be negative. In Table 3 we display the computed values of $c_p(P_{mu})$ obtained with the same test case addressed before. Notice that $c_p(P_{mu}) < 0$ whenever the fine grid becomes small enough, even if GMRES applied to the preconditioned systems does not stagnate and in fact converges in few iterations. As a consequence, the results in Eisenstat *et al.* [21] cannot be directly employed to theoretically justify the convergence observed in our numerical experiments (see Sect. 6).

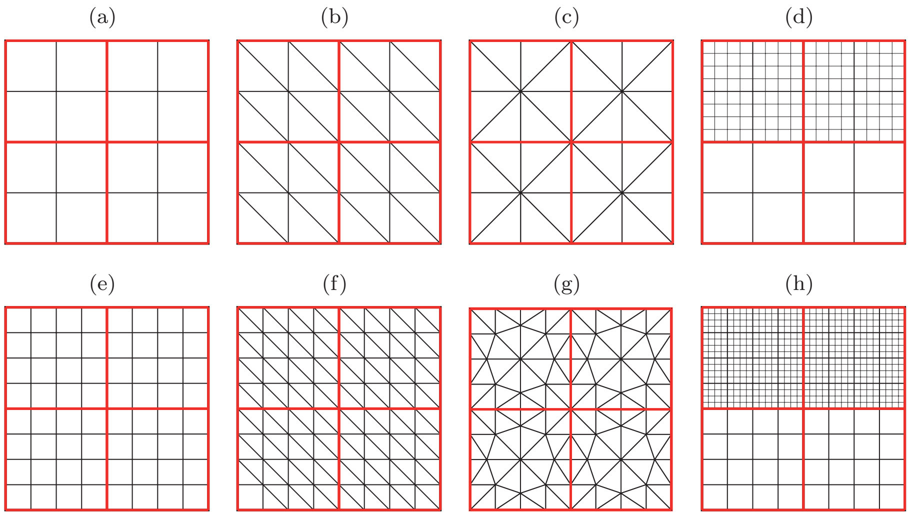

FIGURE 1. Initial coarse (top) and fine (bottom) refinements on Cartesian grids, structured triangular grids, unstructured triangular grids and non-matching Cartesian grids, respectively, on a 4 subdomains partition.

6. Numerical results

In this section we present some two-dimensional numerical experiments to demonstrate the theoretical results of the previous sections.

The subdomain partitions of the domain $\Omega = (0,1) \times (0,1)$ consist of N_s non-overlapping squares, with $N_s = 4, 16$ (see Fig. 1 where the case $N_s = 4$ is shown). We have tested our Schwarz methods on matching and non-matching Cartesian grids, and on structured and unstructured triangular grids. The initial coarse and fine refinements for all the considered triangulations are depicted in Figure 1 (top and bottom, respectively). We have denoted by H_0 and h_0 the corresponding initial coarse and fine mesh sizes. We have considered m successive global uniform refinements of these initial grids so that the resulting mesh sizes are $H_m = H_0/2^m$ and $h_m = h_0/2^m$, respectively, with $m = 0, 1, 2, 3$. All experiments have been carried out with different polynomial approximation degrees for both the fine mesh space ($\ell_h = 1, 2$) and the coarse mesh space ($\ell_H = 0, 1, 2$). For the sake of simplicity, in all the tests we have taken $\alpha_e = \alpha$ for all $e \in \mathscr{E}$. To solve the linear systems, we have used either the (non restarted) GMRES or the CG iterative solvers with (relative) tolerance set to 10^{-9}, allowing for a maximum of 300 iterations (for the non-preconditioned systems we have admitted at most 800 iterations). All computations have been performed in MATLAB.

In Sections 6.1 and 6.2 we investigate the performance of the proposed Schwarz preconditioners for symmetric and non-symmetric DG approximations, respectively. In both sections, f is chosen so that the exact solution of the model problem (2.1) (with non-homogeneous boundary conditions) is given by $u(x, y) = \exp(xy)$. In Section 6.3, we report some numerical experiments carried out with a more general second order elliptic equation with a discontinuous diffusion matrix and well as anisotropic diffusion matrix.

6.1. **Symmetric DG approximations**

We first address the scalability of the proposed multiplicative Schwarz method, that is the independence of the convergence rates on the number of subdomains. In Table 4 we have reported the GMRES iteration counts for the SIPG method ($\alpha = 10$) on the two different subdomain partitions ($N_s = 4, 16$, respectively) by using piecewise bilinear polynomials both for the fine and coarse mesh spaces ($\ell_h = \ell_H = 1$) on structured

TABLE 4. $\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 10$): $\ell_h = \ell_H = 1$, structured triangular grids.

	$N = 4$				$N = 16$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	10	15	21	29	10	14	20	27
$H_0/2$	–	9	13	18	–	9	12	18
$H_0/4$	–	–	8	11	–	–	7	11
$H_0/8$	–	–	–	7	–	–	–	6
# iter($\boldsymbol{A}$)	96	189	368	x	96	189	368	x

TABLE 5. $\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 2$ and $\alpha = 4$): $\ell_h = \ell_H = 1$, Cartesian grids.

	$\alpha = 2$				$\alpha = 4$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	22	27	37	51	22	31	44	61
$H_0/2$	–	21	26	35	–	21	29	41
$H_0/4$	–	–	21	26	–	–	21	29
$H_0/8$	–	–	–	20	–	–	–	20
# iter($\boldsymbol{A}$)	42	75	145	282	53	103	199	388

triangular grids (see Figs. 1(b)–1(f)). The dashes indicate that $\mathcal{T}_H \subseteq \mathcal{T}_h$ and therefore it is meaningless to build the preconditioner. The crosses in Table 4 (and in the rest of the section) indicate that we were not able to solve the non-preconditioned system due to excessive GMRES memory storage requirements. As predicted by Theorem 4.9, our preconditioner is practically insensitive to the number of the subdomains, and, by refining both the fine and the coarse meshes while keeping the ratio H/h constant (the entries in the diagonals of the tables), the iteration counts remain almost unchanged. It can be also seen that, for fixed H we observe an asymptotic behaviour $\mathcal{O}(1/\sqrt{h})$, and for fixed h, the computed convergence behaviour is slightly better than $\mathcal{O}(\sqrt{H})$. In all our computations we have observed that, when $h_0/H_0 = 2$, that is when $H = h$ (see Fig. 1), the coarse solver turns out to be an exact solver for the whole problem, and only one iteration is needed for convergence. Since this is a purely academic case, it is not reported here.

Having already discussed the issue of scalability of the proposed preconditioner, the rest of the numerical experiments of this section have been carried out on a $N_s = 16$ subdomain partition. In Table 5 we report the GMRES iteration counts obtained with the SIPG method ($\alpha = 2$ and $\alpha = 4$, respectively) on Cartesian grids and by choosing $\ell_h = \ell_H = 1$. Our computations indicate that the minimum value $\widetilde{\alpha}$ (see Tab. 1) that guarantees the stability of the scheme is approximately 1.4; therefore, $\alpha = 2$ is very close to the limit case. The results in Table 5 confirm that our preconditioner performs according to the theory also in these cases.

Next, we address the performance of the multiplicative preconditioner for DG methods with higher order polynomial approximation degrees. In Table 6 we have reported the GMRES iteration counts for the LDG method ($\alpha = 1$, $\boldsymbol{\beta} = [0.5, 0.5]^T$) with $\ell_h = \ell_H = 2$, and with $\ell_h = 2$ and $\ell_H = 1$, respectively, on Cartesian grids. Notice that, by choosing $\ell_h = \ell_H = 2$ our preconditioner performs better than with $\ell_h = 2$ and $\ell_H = 1$. The results reported in Table 6 show that, our preconditioner performs well also for small values of the penalty parameter α, confirming that the hypothesis on the size of α required by Theorem 4.9 is only technical and it is not needed in practice. * as already observed in the previous experiments, for fixed H,

TABLE 6. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. LDG method ($\alpha = 1$, $\boldsymbol{\beta} = [0.5, 0.5]^T$): $\ell_h = 2$, $N_s = 16$, Cartesian grids.

	$\ell_H = 2$				$\ell_H = 1$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	12	16	21	28	22	30	40	53
$H_0/2$	–	9	12	16	–	17	23	32
$H_0/4$	–	–	7	8	–	–	16	21
$H_0/8$	–	–	–	5	–	–	–	13
# iter($\boldsymbol{A}$)	112	210	403	x	112	210	403	x

TABLE 7. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 10$): $\ell_H = 0$, $N_s = 16$, structured triangular grids.

	$\ell_h = 1$				$\ell_h = 2$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	25	35	47	66	45	63	87	116
$H_0/2$	–	28	38	53	–	50	70	95
$H_0/4$	–	–	30	41	–	–	54	74
$H_0/8$	–	–	–	30	–	–	–	45
# iter($\boldsymbol{A}$)	96	189	368	x	185	357	688	x

the observed convergence behaviour seems to be of order $\mathcal{O}(1/\sqrt{h})$, and, for fixed h, it seems to be slightly better than $\mathcal{O}(\sqrt{H})$.

Now we investigate the effect of the choice of a piecewise constant coarse mesh space V_H ($\ell_H = 0$). In Table 7 we have reported the results obtained with the SIPG method ($\alpha = 10$) on structured triangular grids with $\ell_h = 1$ and $\ell_h = 2$. For fixed H, a convergence behaviour of order $\mathcal{O}(\sqrt{h^{-1}})$ is clearly observed, while, for fixed h, a convergence behaviour of order $\mathcal{O}(\sqrt{H})$ seems to be achieved only asymptotically. Such a behaviour is similar to that observed in [1,22] for the additive Schwarz preconditioner (3.7) with a piecewise constant coarse solver (see also Tab. 9, below). Results carried out with the LDG method on Cartesian grids (not reported here, for the sake of brevity), confirm the convergence behaviour observed with the SIPG method.

Next, we present a comparison between the multiplicative Schwarz method (3.6) and its additive version (3.7) introduced in [1]. As we have already noticed, for symmetric DG methods the additive Schwarz preconditioner (3.7) is symmetric and therefore, the resulting preconditioned linear systems could be solved with the CG method which is considerably cheaper (in terms of computational costs) than the GMRES iterative solver. However, in order to present a fair comparison, we have solved all the preconditioned linear systems with the GMRES iterative solver. The following numerical experiments have been carried out on non-matching grids (see Figs. 1(d)–1(h)). In Table 8 we compare the iteration counts obtained with $\ell_h = \ell_H = 1$; the analogous results carried out with $\ell_h = 1$ and $\ell_H = 0$ are shown in Table 9. In both cases, the multiplicative Schwarz method is far faster than the additive one and, when using $\ell_h = \ell_H = 1$, the observed convergence behaviour is in agreement with the theoretical results given Theorem 4.9. Concerning the results obtained with $\ell_h = 1$ and $\ell_H = 0$, for fixed H, a convergence behaviour of order $\mathcal{O}(1/\sqrt{h})$ is clearly achieved with both preconditioners, while, for fixed H, the multiplicative Schwarz preconditioner seems to perform slightly better than the additive one.

TABLE 8. $\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$ and $\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 10$): $\ell_h = \ell_H = 1$, $N_s = 16$, non-matching Cartesian grids.

	$\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$				$\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$			
$H \downarrow \; h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	13	20	30	42	43	64	94	133
$H_0/2$	–	10	16	24	–	47	68	97
$H_0/4$	–	–	9	14	–	–	48	69
$H_0/8$	–	–	–	8	–	–	–	48
# iter($\boldsymbol{A}$)	277	538	x	x	277	538	x	x

TABLE 9. $\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$ and $\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 10$): $\ell_h = 1$, $\ell_H = 0$, $N_s = 16$, non-matching Cartesian grids.

	$\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$				$\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$			
$H \downarrow \; h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	45	64	89	122	90	128	182	264
$H_0/2$	–	47	66	91	–	103	149	215
$H_0/4$	–	–	46	63	–	–	119	174
$H_0/8$	–	–	–	44	–	–	–	134
# iter($\boldsymbol{A}$)	277	538	x	x	277	538	x	x

TABLE 10. $\mathbf{B}_{mu}^{sym}\mathbf{Au} = \mathbf{B}_{mu}^{sym}\mathbf{f}$: condition number estimates and CG iteration counts. SIPG method ($\alpha = 10$), $\ell_h = \ell_H = 1$, Cartesian grids.

	$\kappa(\mathbf{B}_{mu}^{sym}\mathbf{A})$					CG # iter($\mathbf{B}_{mu}^{sym}\mathbf{A}$)			
$H \downarrow \; h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$		h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	5.2	10.4	21.2	43.1		13	21	29	37
$H_0/2$	–	4.8	9.4	18.7		–	11	17	23
$H_0/4$	–	–	4.7	9.4		–	–	10	15
$H_0/8$	–	–	–	4.5		–	–	–	9
$\kappa(\boldsymbol{A})$	2.7 e+02	10 e+03	4.2 e+03	1.7 e+04	# iter($\boldsymbol{A}$)	77	155	303	592

Now, we present a set of numerical experiments carried out with the symmetrized version of the multiplicative Schwarz operator P_{mu}^{sym} defined in (3.8). In Table 10 we have reported the condition number estimates and CG iteration counts computed with the SIPG method ($\alpha = 10$) by choosing $\ell_h = \ell_H = 1$ on Cartesian grids. The numerical estimate of the condition number has been obtained by exploiting the analogies between the Lanczos technique and the CG method; in fact, a tridiagonal matrix can be built in the CG code with the property that, during the iterative procedure, the approximation of the extreme eigenvalues of $\mathbf{B}_{mu}^{sym}\mathbf{A}$ becomes better and better (see [23], Sects. 9.3 and 10.2 for more details). Clearly, the numerical results reported in Table 10 confirm the theoretical condition number estimate given in (4.17). These results can be compared with the ones given in [1], Table 9, and show that $\kappa(P_{mu}^{sym})$ is much smaller than $\kappa(P_{ad})$.

Now we present some results obtained with the BRMPS method [6] which is a DG method stabilized by means of the operator $\mathcal{S}^r(\cdot,\cdot)$ (see (2.6)) and to which our analysis does not apply. In Table 11 we show the GMRES iteration counts obtained with the BRMPS method ($\alpha = 10$) on Cartesian grids with $\ell_h = 1, 2$ and $\ell_H = 1$.

TABLE 11. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. BRMPS method ($\alpha = 10$): $N_s = 16$, Cartesian grids.

	$\ell_h = 1$, $\ell_H = 1$				$\ell_h = 2$, $\ell_H = 1$			
$H \downarrow$ $h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	18	28	38	54	28	40	54	73
$H_0/2$	–	15	23	32	–	23	33	44
$H_0/4$	–	–	12	20	–	–	20	28
$H_0/8$	–	–	–	11	–	–	–	17
# iter($\boldsymbol{A}$)	83	167	325	630	161	313	602	x

TABLE 12. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. NIPG method ($\alpha = 1$): $\ell_h = \ell_H = 1$, Cartesian grids.

	$N = 4$				$N = 16$			
$H \downarrow$ $h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	9	12	15	18	12	13	16	20
$H_0/2$	–	9	11	13	–	9	11	14
$H_0/4$	–	–	8	10	–	–	8	10
$H_0/8$	–	–	–	7	–	–	–	7
# iter($\boldsymbol{A}$)	33	61	117	225	33	61	117	225

In both cases, the multiplicative preconditioner seems to perform very well, and the observed convergence behaviour is similar to the one observed with all the DG methods covered by our theoretical analysis.

6.2. Non-symmetric DG approximations

This section is devoted to show some numerical experiments carried out with the proposed multiplicative preconditioner for non-symmetric DG approximations of the model problem (2.1). We shall show that, although the available theory of Schwarz methods and the GMRES convergence theory of Eisenstat *et al.* [21] cannot be used to explain the convergence of the proposed preconditioners, they do indeed converge when accelerated with the GMRES and they are also scalable.

As in the previous section, we first address the scalability of the proposed multiplicative preconditioner. In Table 12 we have reported the iterations counts obtained on Cartesian grids with the NIPG method ($\alpha = 1$) and with $\ell_h = \ell_H = 1$ on a $N_s = 4$ (left) and $N_s = 16$ (right) subdomain partition. It can be observed that the multiplicative preconditioner seems to be scalable.

Throughout the rest of the section, we set $N_s = 16$. Since the NIPG method is stable for any $\alpha > 0$, we have tested our preconditioner with $\alpha = 0.1, 0.2$. The results obtained on Cartesian grids with $\ell_h = \ell_H = 1$ are shown in Table 13. As it can be clearly observed, our preconditioner performs well also by choosing $\alpha < 1$.

Now we discuss the performance of the proposed preconditioner for DG approximations with higher order polynomial approximation degrees. The results reported in Table 14 have been carried out with the NIPG method ($\alpha = 1$) on structured triangular grids. We have set $\ell_h = 2$ and in Table 14 we have compared the GMRES iteration counts obtained with $\ell_H = 2$ and with $\ell_H = 1$. We observe that, in both cases, the number of iterations do not increase as we decrease both h and H keeping their ratio constant. Furthermore, as for symmetric DG approximations (see Sect. 6.1), the proposed preconditioner behaves better with equal order polynomial approximation degree, *i.e.*, $\ell_h = \ell_H = 2$, than with $\ell_h = 2$ and $\ell_H = 1$.

TABLE 13. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. NIPG method ($\alpha = 0.1$ and $\alpha = 0.2$): $\ell_h = \ell_H = 1$, Cartesian grids.

	$\alpha = 0.1$				$\alpha = 0.2$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	57	59	57	63	48	46	46	53
$H_0/2$	–	47	46	50	–	35	37	42
$H_0/4$	–	–	41	47	–	–	32	37
$H_0/8$	–	–	–	40	–	–	–	31
# iter($\boldsymbol{A}$)	40	60	107	208	36	57	109	211

TABLE 14. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. NIPG method ($\alpha = 1$): $\ell_h = 2$, $N_s = 16$, structured triangular grids.

	$\ell_h = 2,\ \ell_H = 2$				$\ell_h = 2,\ \ell_H = 1$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	11	13	16	17	16	19	23	30
$H_0/2$	–	9	11	13	–	13	16	20
$H_0/4$	–	–	7	9	–	–	12	14
$H_0/8$	–	–	–	5	–	–	–	11
# iter($\boldsymbol{A}$)	119	222	426	x	119	222	426	x

TABLE 15. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. NIPG method ($\alpha = 1$): $N_s = 16$, unstructured triangular grids.

	$\ell_h = 1,\ \ell_H = 0$				$\ell_h = 2,\ \ell_H = 0$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	25	32	43	54	26	37	52	74
$H_0/2$	–	27	34	44	–	30	42	61
$H_0/4$	–	–	26	35	–	–	33	48
$H_0/8$	–	–	–	27	–	–	–	35
# iter($\boldsymbol{A}$)	54	94	192	x	122	225	459	x

Next, we study the effect of using a coarse mesh space V_H made of piecewise constants ($\ell_H = 0$) on the performance of our multiplicative Schwarz method. In Table 15 we compare the results obtained with the NIPG method ($\alpha = 1$) on unstructured triangular grids with $\ell_h = 1$ and $\ell_h = 2$. Similar conclusions as for symmetric DG approximations (see Sect. 6.1) can be drawn: for fixed H, the convergence behaviour seems to be $\mathcal{O}(1/\sqrt{h})$, while, for fixed h, a convergence behaviour of order $\mathcal{O}(\sqrt{H})$ seems to be achieved only asymptotically. Such a behaviour was already observed in [1] with the additive Schwarz preconditioner.

Finally, we present a numerical comparison between our multiplicative Schwarz method and the additive one (3.7) proposed in [1], which, in this case, is non-symmetric as well since we are dealing with non-symmetric DG schemes. Results in Table 16 have been carried out with the NIPG method ($\alpha = 1$) on non-matching Cartesian grids (see Figs. 1(d)–1(h)). More precisely, in Table 16 we compare the GMRES iteration counts for the multiplicative and the additive Schwarz preconditioners, respectively, with $\ell_h = \ell_H = 1$.

TABLE 16. $\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$ and $\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$: GMRES iteration counts. NIPG method ($\alpha = 1$): $\ell_h = \ell_H = 1$, $N_s = 16$, non-matching Cartesian grids.

	$\mathbf{B}_{mu}\mathbf{Au} = \mathbf{B}_{mu}\mathbf{f}$				$\mathbf{B}_{ad}\mathbf{Au} = \mathbf{B}_{ad}\mathbf{f}$			
$H \downarrow\ h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	11	13	17	21	28	30	36	46
$H_0/2$	–	9	12	14	–	26	30	36
$H_0/4$	–	–	8	10	–	–	26	29
$H_0/8$	–	–	–	7	–	–	–	25
# iter($\boldsymbol{A}$)	113	217	x	x	113	217	x	x

(a) (b)

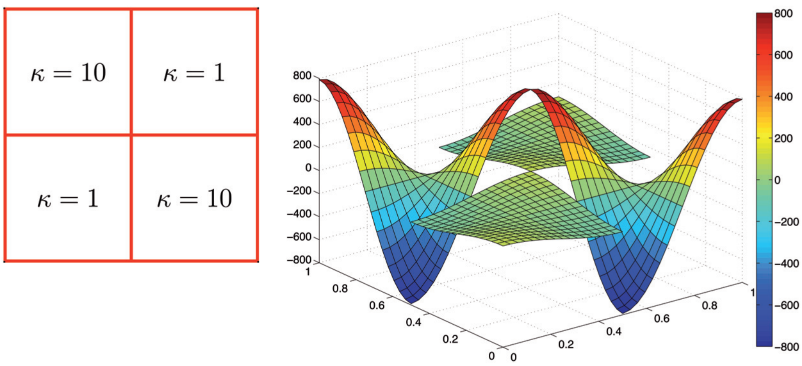

FIGURE 2. Test case with $u(x,y) = \cos(2\pi x)\cos(2\pi y)$. Figure 2(a): the discontinuous permeability coefficient κ. Figure 2(b): the elevation plot of the source term $f(x,y)$.

Clearly, the proposed multiplicative Schwarz method outperforms the additive one, in the sense that it is more than twice faster (recall that, in terms of the computational cost the multiplicative preconditioner is more expensive). We have repeated the same set of experiments with $\ell_H = 0$; the numerical results (not reported here) confirmed that also with a piecewise constant coarse solver the multiplicative Schwarz preconditioner is twice faster than the additive one.

6.3. Discontinuous and anisotropic diffusion matrices

We let $\Omega = (0,1) \times (0,1)$, and we consider the following diffusion problem

$$-\operatorname{div}(\mathtt{K}\,\nabla u) = f \text{ in } \Omega, \quad u = g \text{ on } \partial\Omega,$$

where $\mathtt{K}$ is a *symmetric* 2×2 matrix whose entries are bounded, piecewise continuous, strictly positive coefficients.

In the first test case, we have taken $u(x,y) = \cos(2\pi x)\cos(2\pi y)$, and $\mathtt{K} = \kappa I$, where κ is a piecewise constant positive coefficient with a discontinuity between four regions of Ω (checkerboard domain) chosen as in Figure 2(a); the load f (*cf.* Fig. 2(b)) and the boundary conditions are set accordingly. We have aligned the subdomain partition with the discontinuities of the diffusion coefficient κ as shown in Figure 2(a).

In Table 17 we have reported the GMRES iteration counts obtained on Cartesian grids for the LDG method ($\alpha = 1$, $\boldsymbol{\beta} = [0.5, 0.5]^T$) with $\ell_h = \ell_H = 1$ and with $\ell_h = \ell_H = 2$. As expected, we observe that the presence of a discontinuous diffusion coefficient does not deteriorate the numerical performance of our preconditioner, and the expected convergence rates are attained.

TABLE 17. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. LDG method ($\alpha = 1$, $\boldsymbol{\beta} = [0.5, 0.5]^T$): $\ell_h = 2$, $N_s = 4$, Cartesian grids, κ as in Figure 2(a).

	$\ell_h = 1 = \ell_H = 1$				$\ell_h = 2 = \ell_H = 2$			
$H \downarrow \; h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	24	39	56	81	25	37	54	77
$H_0/2$	–	31	46	61	–	27	35	50
$H_0/4$	–	–	35	47	–	–	25	29
$H_0/8$	–	–	–	36	–	–	–	18
# iter($\boldsymbol{A}$)	112	220	387	x	187	354	643	x

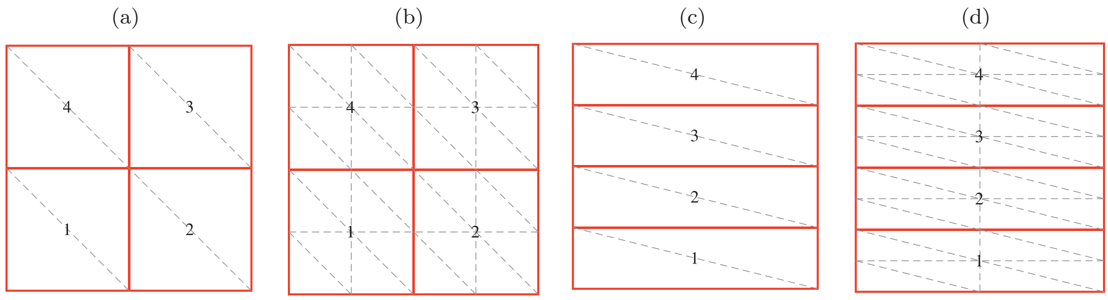

FIGURE 3. The two different subdomain partitions in the test case with $\mathtt{K} = \mathrm{diag}(10^{-2}, 1)$, and the corresponding initial coarse grids (Figs. 3(a)–3(c)) and fine grids (Figs. 3(b)–3(d)).

TABLE 18. $\mathbf{B}_{mu}\mathbf{A}\mathbf{u} = \mathbf{B}_{mu}\mathbf{f}$: GMRES iteration counts. SIPG method ($\alpha = 10$): $\ell_h = \ell_H = 1$, $\mathtt{K} = \mathrm{diag}(10^{-2}, 1)$, structured triangular grids.

	Subdomain partition as in Figures 3(a)–3(b)				Subdomain partition as in Figures 3(c)–3(d)			
$H \downarrow \; h \rightarrow$	h_0	$h_0/2$	$h_0/4$	$h_0/8$	h_0	$h_0/2$	$h_0/4$	$h_0/8$
H_0	15	28	50	82	1	1	1	1
$H_0/2$	–	20	35	56	–	1	1	1
$H_0/4$	–	–	22	39	–	–	1	1
$H_0/8$	–	–	–	24	–	–	–	1
# iter($\boldsymbol{A}$)	73	180	409	x	77	189	416	x

In the second test case, the exact solution is chosen as $u(x, y) = \exp(xy)$, and $\mathtt{K} = \mathrm{diag}(10^{-2}, 1)$. Here the problem exhibits a strong anisotropy along the x-direction. In Table 18 we have reported the GMRES iteration counts obtained with the SIPG method ($\alpha = 10$), $\ell_h = \ell_H = 1$ and with the two different subdomain partitions shown in Figure 3. More precisely, in Table 18 (left) we show the iterations counts obtained with the subdomain partition shown in Figures 3(a)–3(b), whereas the analogous results obtained with a subdomain partition aligned with the anisotropy of the problem are reported in Table 18 (right). For the former subdomain partition we observe the expected convergence rates, whereas for the latter only one iteration is needed to achieve the required precision.

7. Conclusions

We have introduced and analyzed some new non-overlapping multiplicative Schwarz methods for DG approximations of second order elliptic operators in divergence form. The construction of the preconditioners is presented in a unified framework, allowing also for the use of non-matching grids. Due to the lack of symmetry of some of the DG methods considered, the analysis is developed by distinguishing between symmetric and non-symmetric DG approximations. For a wide class of symmetric DG methods we have proved, under a technical hypothesis on the size of the penalty parameter, optimal convergence estimates for the resulting iterative methods accelerated with suitable Krylov space solvers. We have addressed the issue of preconditioning non-symmetric DG approximations, demonstrating numerically that neither the abstract Schwarz convergence theory given [15] nor the GMRES convergence theory of Eisenstat *et al.* [21] can be applied to explain the convergence observed in the numerical experiments. The issue of exploring other GMRES convergence theories to explain theoretically the optimal behaviour observed is under investigation. Extensive numerical experiments confirm the theoretical results and assess the performance of the proposed preconditioners. A numerical comparison between the proposed multiplicative preconditioner and the additive one studied in [1] have been also included.

Acknowledgements. The authors are grateful to Annalisa Buffa and Ilaria Perugia for the multiple discussions which took place while developing this work and for the careful reading of the manuscript. Special thanks go to Valeria Simoncini for her precious guide on linear algebra topics. The authors are also extremely thankful to Susan Brenner and Li-Yeng Sung for the constructive discussions on the issue of preconditioning non-symmetric DG methods that took place when they were visiting Pavia.

References

[1] P.F. Antonietti and B. Ayuso, Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: Non-overlapping case. *ESAIM: M2AN* **41** (2007) 21–54.

[2] P.F. Antonietti, A. Buffa and I. Perugia, Discontinuous Galerkin approximation of the Laplace eigenproblem. *Comput. Methods Appl. Mech. Engrg.* **195** (2006) 3483–3503.

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

[4] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39** (2001) 1749–1779 (electronic).

[5] F. Bassi and S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.* **131** (1997) 267–279.

[6] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti and M. Savini, A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows, in *Proceedings of the 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics*, R. Decuypere and G. Dibelius Eds., Technologisch Instituut, Antwerpen, Belgium (1997) 99–108.

[7] C.E. Baumann and J.T. Oden, A discontinuous *hp* finite element method for convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.* **175** (1999) 311–341.

[8] J.H. Bramble, J.E. Pasciak, J.P. Wang and J. Xu, Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.* **57** (1991) 1–21.

[9] S.C. Brenner and O. Luke, A W-cycle algorithm for a weakly over-penalized interior penalty method. *JNAIAM J. Numer. Anal. Indust. Appl. Math* **196** (2007) 3823–3832.

[10] S.C. Brenner and O. Luke, A weakly over-penalized non-symmetric Interior Penalty method. *Comput. Methods Appl. Mech. Engrg.* **2** (2007) 35–48.

[11] S.C. Brenner and L.-Y. Sung, Multigrid algorithms for C^0 interior penalty methods. *SIAM J. Numer. Anal.* **44** (2006) 199–223 (electronic).

[12] S.C. Brenner and K. Wang, Two-level additive Schwarz preconditioners for C^0 interior penalty methods. *Numer. Math.* **102** (2005) 231–255.

[13] S.C. Brenner and J. Zhao, Convergence of multigrid algorithms for interior penalty methods. *Appl. Numer. Anal. Comput. Math.* **2** (2005) 3–18.

[14] F. Brezzi, G. Manzini, D. Marini, P. Pietra and A. Russo, Discontinuous Galerkin approximations for elliptic problems. *Numer. Methods Partial Differential Equations* **16** (2000) 365–378.

[15] X.-C. Cai and O.B. Widlund, Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM J. Numer. Anal.* **30** (1993) 936–952.

[16] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, *Studies in Mathematics and its Applications* **4**. North-Holland Publishing Co., Amsterdam (1978).

[17] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.* **35** (1998) 2440–2463 (electronic).

[18] C. Dawson, S. Sun and M.F. Wheeler, Compatible algorithms for coupled flow and transport. *Comput. Methods Appl. Mech. Engrg.* **193** (2004) 2565–2580.

[19] V.A. Dobrev, R.D. Lazarov, P.S. Vassilevski and L.T. Zikatanov, Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. *Numer. Linear Algebra Appl.* **13** (2006) 753–770.

[20] J. Douglas, Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, in *Computing Methods in Applied Sciences (Second Internat. Sympos., Versailles, 1975)*, *Lecture Notes in Physics* **58**, Springer, Berlin (1976) 207–216.

[21] S.C. Eisenstat, H.C. Elman and M.H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.* **20** (1983) 345–357.

[22] X. Feng and O.A. Karakashian, Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.* **39** (2001) 1343–1365 (electronic).

[23] G.H. Golub and C.F. Van Loan, *Matrix Computations*. 3rd Edn., Johns Hopkins University Press, Baltimore, USA (1996).

[24] J. Gopalakrishnan and G. Kanschat, A multilevel discontinuous Galerkin method. *Numer. Math.* **95** (2003) 527–550.

[25] G. Kanschat, Preconditioning methods for local discontinuous Galerkin discretizations. *SIAM J. Sci. Comput.* **25** (2003) 815–831 (electronic).

[26] G. Kanschat, Block preconditioners for LDG discretizations of linear incompressible flow problems. *J. Sci. Comput.* **22/23** (2005) 371–384.

[27] C. Lasser and A. Toselli, An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.* **72** (2003) 1215–1238 (electronic).

[28] P.-L. Lions, On the Schwarz alternating method. I, in *First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987)*, SIAM, Philadelphia, PA (1988) 1–42.

[29] M. Murillo and X.-C. Cai, A fully implicit parallel algorithm for simulating the non-linear electrical activity of the heart. *Numer. Linear Algebra Appl.* **11** (2004) 261–277.

[30] L.F. Pavarino and S. Scacchi, Multilevel Schwarz and multigrid preconditioners for the bidomain system, in *Domain Decomposition Methods in Science and Engineering* **XVII**, U. Langer, M. Discacciati, D. Keyes, O. Widlund and W. Zulehner Eds., *Lecture Notes in Computational Science and Engineering* **60**, Springer, Heidelberg (2008) 631–638.

[31] L.F. Pavarino and A. Toselli, *Recent Developments in Domain Decomposition Methods*, *Lecture Notes in Computational Science and Engineering* **23**. Springer-Verlag, Berlin (2002). [Selected papers from the Workshop on Domain Decomposition held at ETH Zürich, Zürich, June 7–8 (2001)].

[32] W.H. Reed and T. Hill, *Triangular mesh methods for the neutron transport equation*. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, USA (1973).

[33] B. Rivière, M.F. Wheeler and V. Girault, Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I. *Comput. Geosci.* **3** (1999) 337–360.

[34] V. Simoncini and D.B. Szyld, *New conditions for non-stagnation of minimal residual methods*. Technical Report 07-04-17, Department of Mathematics, Temple University, USA (2007), to appear in *Numerische Mathematik*.

[35] B.F. Smith, P.E. Bjørstad and W.D. Gropp, *Domain decomposition. Parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press, Cambridge (1996).

[36] R. Stenberg, Mortaring by a method of J.A. Nitsche, in *Computational mechanics (Buenos Aires, 1998)*, Centro Internac. Métodos Numér. Ing., Barcelona, Spain (1998).

[37] A. Toselli and O. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, *Springer Series in Computational Mathematics* **34**. Springer-Verlag, Berlin (2005).

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

[39] J. Xu, Iterative methods by SPD and small subspace solvers for nonsymmetric or indefinite problems, in *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations (Norfolk, VA, 1991)*, SIAM, Philadelphia, PA (1992) 106–118.

[40] J. Xu, A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM J. Numer. Anal.* **29** (1992) 303–319.

[41] J. Xu and L. Zikatanov, The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Amer. Math. Soc.* **15** (2002) 573–597 (electronic).