SIAM J. NUMER. ANAL.
Vol. 53, No. 1, pp. 573–597

ANALYSIS OF SCHWARZ METHODS FOR A HYBRIDIZABLE DISCONTINUOUS GALERKIN DISCRETIZATION*

MARTIN J. GANDER† AND SOHEIL HAJIAN†

Abstract. Schwarz methods are attractive parallel solvers for large-scale linear systems obtained when partial differential equations are discretized. For hybridizable discontinuous Galerkin (HDG) methods, this is a relatively new field of research, because HDG methods impose continuity across elements using a Robin condition, while classical Schwarz solvers use Dirichlet transmission conditions. Robin conditions are used in optimized Schwarz methods to get faster convergence compared to classical Schwarz methods, and this even without overlap, when the Robin parameter is well chosen. We present in this paper a rigorous convergence analysis of Schwarz methods for the concrete case of the hybridizable interior penalty (IPH) method. We show that the penalization parameter needed for convergence of IPH leads to slow convergence of the classical additive Schwarz method, and we propose a modified solver which leads to much faster convergence. Our analysis is entirely at the discrete level and thus holds for arbitrary interfaces between two subdomains. We then generalize the method to the case of many subdomains, including cross-points, and obtain a new class of preconditioners for Krylov subspace methods which exhibit better convergence properties than the classical additive Schwarz preconditioner. We illustrate our results with numerical experiments.

Key words. additive Schwarz, optimized Schwarz, discontinuous Galerkin methods

AMS subject classifications. 65N22, 65F10, 65F08, 65N55, 65H10

DOI. 10.1137/140961857

1. Introduction.

We consider the elliptic model problem

$$\begin{aligned} \eta(x)u(x) - \nabla\cdot(a(x)\nabla u) &= f && \text{in } \Omega\subset\mathbb{R}^2,\\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

in the weak sense where $f \in \mathrm{L}^2(\Omega)$, $a(x) \in \mathrm{L}^\infty(\Omega)$ and uniformly positive, $\eta_0 \geq \eta(x) \geq 0$, and Ω is assumed to be a convex polygon for simplicity. Any discretization of this problem, for example, by a finite element method (FEM) or a discontinuous Galerkin (DG) method, leads to a large sparse linear system

$$A\boldsymbol{u} = \boldsymbol{f}, \tag{1.2}$$

where $\boldsymbol{u}$ is the vector of degrees of freedom (DOFs) representing an approximation of u and A represents the disretized differential operator. In this paper we consider a hybridizable interior penalty (IPH[1]) discretization which results in a symmetric positive definite (s.p.d.) matrix A. An IPH discretization seeks $u_h \in \mathrm{L}^2(\Omega)$ over a triangulation of the domain where u_h is not necessarily continuous across elements. As common to DG methods, IPH imposes the continuity of the solution approximately through penalization techniques, i.e., penalizing jumps of u_h across elements in the bilinear form. The penalization is controlled by a penalty parameter μ.

Since the matrix A of IPH is s.p.d. and sparse, one can use the conjugate gradient (CG) method to solve the linear system (1.2). The convergence of CG slows down as

*Received by the editors March 21, 2014; accepted for publication (in revised form) December 9, 2014; published electronically February 19, 2015.
http://www.siam.org/journals/sinum/53-1/96185.html

†Section de Mathematiques, University of Geneva, Geneva CH-1211, Switzerland (martin.gander@unige.ch, soheil.hajian@unige.ch).

[1] We use the acronym IPH for *hybridizable interior penalty* because this has become the common abbreviation following its introduction in [7] as a member of the family of HDG methods.

the condition number $\kappa(A)$ grows. It is not hard to show that $\kappa(A) = O(h^{-2})$, where h is the maximum diameter of the elements in the triangulation; see, for instance, [6]. Therefore preconditioning is unavoidable and domain decomposition preconditioners have been developed and studied for such discretizations; see [2, 12]. IPH as local solvers were also used to precondition classical IP discretizations [1]. One can also design a substructuring preconditioner for a p-version of IPH with polylogarithmic growth in the condition number; see for details [24]. For a similar discretization where the approximation is continuous inside subdomains but discontinuous across subdomains, a substructuring preconditioner was proposed and analyzed for the h-version with logarithmic growth in the condition number; see [9].

A favorite preconditioner is the additive Schwarz preconditioner, for which the set of unknowns is partitioned into overlapping or nonoverlapping subsets, corresponding to subdomains with maximum diameter H. In this paper we consider only the nonoverlapping case[2] and for simplicity study first only two subdomains; a generalization is given in section 5. The nonoverlapping two-subdomain decomposition results in a natural partitioning of the unknowns $\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2)^\top$. The solution of the linear system by the additive Schwarz method without overlap is equivalent to the block Jacobi iteration

$$M\mathbf{u}^{(n+1)} = N\mathbf{u}^{(n)} + \mathbf{f}, \quad M = \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}, \ N = M - A. \tag{1.3}$$

The matrix M is also s.p.d. and can be considered as a preconditioner for CG. It can be shown that in this case we have $\kappa(M^{-1}A) \leq O(h^{-1})$ in the absence of a coarse solver; see [12]. Preconditioned CG then satisfies the convergence factor estimate $\rho \leq \frac{\sqrt{\kappa(M^{-1}A)}-1}{\sqrt{\kappa(M^{-1}A)}+1} = 1 - O(\sqrt{h})$.

On the other hand it has been recently shown in [15] that the block Jacobi iteration in (1.3) for an IPH discretization can be viewed as a discretization of a nonoverlapping Schwarz method with Robin transmission conditions, i.e.,

$$\begin{array}{rclll} (\eta - \Delta)u_1^{(n+1)} & = & f & \text{in } \Omega_1, & (\eta - \Delta)u_2^{(n+1)} = f \quad \text{in } \Omega_2, \\ \mathcal{B}_1 u_1^{(n+1)} & = & \mathcal{B}_1 u_2^{(n)} & \text{on } \Gamma, & \mathcal{B}_2 u_2^{(n+1)} = \mathcal{B}_2 u_1^{(n)} \quad \text{on } \Gamma, \end{array} \tag{1.4}$$

where $\mathcal{B}_i w = \mu\, w + \frac{\partial w}{\partial \boldsymbol{n}_i}$, Γ is the interface between the two subdomains, and μ is precisely the penalty parameter of the IPH discretization. This parameter μ has to be chosen such that it ensures coercivity and optimal approximation properties. For an IPH discretization, we must have $\mu = \alpha h^{-1}$ for some constant $\alpha > 0$ large enough, independent of h, and this scaling cannot be weakened, since otherwise coercivity is lost. On the other hand, optimized Schwarz theory suggests that the iteration in (1.4) converges faster if $\mu = O(h^{-1/2})$; see [13]. In that case for the contraction factor we have $\rho = 1 - O(\sqrt{h})$, while with the choice $\mu = O(h^{-1})$ for IPH, we have $\rho = 1 - O(h)$.

The challenge is therefore to design a Schwarz algorithm for IPH with convergence factor $\rho = 1 - O(\sqrt{h})$, while having the same fixed point as the original additive Schwarz or block Jacobi method for IPH. An idea for doing this can be found for

[2]There is a subtle difference between overlap at the continuous level of the subdomains and the discrete level of unknowns (see [14]): no overlap at the level of unknowns means minimal overlap of one mesh size at the continuous level for classical discretizations like finite elements or finite differences. This becomes, however, even more subtle here with DG discretizations, since the discrete unknowns are coupled through Robin conditions, and no overlap at the level of unknowns really means no overlap at the continuous level; see [15].

Maxwell's equation in [10]. This approach was also adopted for IPH in [20], where numerical experiments show that the convergence factor is indeed $\rho = 1 - O(\sqrt{h})$, while maintaining the same fixed point, but there is no convergence analysis.

We provide in this paper a convergence theory for Schwarz methods applied to IPH discretizations and prove these numerical observations. A similar analysis exists for classical FEM using Schur complement formulations and exploiting eigenvalues of the Dirichlet-to-Neumann (DtN) operator; see [22]. Our analysis uses similar DtN arguments but is substantially different from [22], since in a DG method continuity conditions are imposed only weakly. We focus in our analysis on the h-version with polynomial degree one and do not study the effect of possible jumps in $a(x)$ or higher polynomial degree.

Our paper is organized as follows: in section 2 we describe two different but equivalent formulations of IPH and construct a Schur complement system. In section 3 we provide mathematical tools to analyze Schwarz methods formulated using Schur complements. In section 4 we present the additive Schwarz and a new Schwarz algorithm for IPH in a two-subdomain setting and prove their convergence with concrete contraction factor estimates. Section 5 contains a generalization of the algorithms to the multi-subdomain case. We show numerical experiments in section 6 to illustrate our analysis and also verify numerically that the new algorithm provides a better preconditioner for Krylov subspace methods: we observe that the contraction factor is $\rho = 1 - O(h^{1/4})$, which is much faster than the CG solver preconditioned by one level additive Schwarz.

2. IPH method. This section is devoted to recalling the definition of IPH in two different but equivalent forms, namely, primal and hybridizable formulation. In section 4 we design and analyze two Schwarz methods for the hybridizable form and show that the first one is slow and equivalent to a block Jacobi method applied to a primal form, i.e., (1.3). However, the second Schwarz method takes advantage of hybridizable formulation and achieve faster convergence.

IPH was first introduced in [11] as a stabilized discontinuous FEM and later was studied as a member of the class of hybridizable DG methods in [7]. It has been shown that it is equivalent to a method called the ultra weak variational formulation for the Helmholtz equation; see [19]. IPH also fits into the framework developed in [3] for a unified analysis of DG methods. IPH is further studied in [21] in the context of incompressible flows.

2.1. Notation. We follow the notation introduced in [3]. Let $\mathcal{T}_h = \{K\}$ be a shape-regular and quasi-uniform triangulation of the domain Ω. Let h_K be the diameter of an element of the triangulation defined by $h_K := \max_{x,y \in K} |x - y|$ and $h = \max_{K \in \mathcal{T}_h} h_K$. If e is an edge of an element, we denote by h_e the length of that edge. The quasi-uniformity of the mesh implies $h \approx h_K \approx h_e$.

We denote by $\mathcal{E}^0$ the set of interior edges shared by two elements in $\mathcal{T}_h$, that is,

$$\mathcal{E}^0 := \{e = \partial K_1 \cap \partial K_2 \ \forall K_1, K_2 \in \mathcal{T}_h\},$$

by $\mathcal{E}^\partial$ the set of boundary edges, and by $\mathcal{E} := \mathcal{E}^\partial \cup \mathcal{E}^0$ all edges. We introduce the broken Sobolev space $\mathrm{H}^l(\mathcal{T}_h) := \prod_{K \in \mathcal{T}_h} \mathrm{H}^l(K)$, where $\mathrm{H}^l(K)$ is the Sobolev space in $K \in \mathcal{T}_h$ and l is a positive integer. Note that $q \in \mathrm{H}^l(\mathcal{T}_h)$ is not necessarily continuous across elements. Therefore the element boundary traces of functions in $\mathrm{H}^l(\mathcal{T}_h)$ belong to $\mathrm{T}(\mathcal{E}) = \prod_{K \in \mathcal{T}_h} \mathrm{L}^2(\partial K)$, where $q \in \mathrm{T}(\mathcal{E})$ can be double-valued on $\mathcal{E}^0$ but is single-valued on $\mathcal{E}^\partial$.

We now define two trace operators: let $q \in \mathrm{T}(\mathcal{E})$ and $q_i := q|_{\partial K_i}$. Then on $e = \partial K_1 \cap \partial K_2$ we define the average and jump operators

$$\{\!\{q\}\!\} := \tfrac{1}{2}(q_1 + q_2), \qquad [\![q]\!] := q_1\,\boldsymbol{n}_1 + q_2\,\boldsymbol{n}_2,$$

where $\boldsymbol{n}_i$ is the unit outward normal from K_i on $e \in \mathcal{E}^0$. It is clear that these operators are independent of the element enumeration. Similarly for a vector-valued function $\boldsymbol{\sigma} \in [\mathrm{T}(\mathcal{E})]^2$ we define on interior edges

$$\{\!\{\boldsymbol{\sigma}\}\!\} := \tfrac{1}{2}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2), \qquad [\![\boldsymbol{\sigma}]\!] := \boldsymbol{\sigma}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{\sigma}_2 \cdot \boldsymbol{n}_2.$$

On the boundary, we set the average and jump operators to $\{\!\{\boldsymbol{\sigma}\}\!\} := \boldsymbol{\sigma}$ and $[\![q]\!] = q\,\boldsymbol{n}$. We do not need to define $\{\!\{q\}\!\}$ and $[\![\boldsymbol{\sigma}]\!]$ on $e \in \mathcal{E}^\partial$.

We define a finite dimensional subspace of $\mathrm{H}^l(\mathcal{T}_h)$ by

$$V_h := \left\{ v \in \mathrm{L}^2(\Omega) : v|_K \in \mathbb{P}^k(K)\ \forall K \in \mathcal{T}_h \right\}, \tag{2.1}$$

where $\mathbb{P}^k(K)$ is the space of polynomials of degree $\le k$ in the simplex $K \in \mathcal{T}_h$. We denote boundary integrals on an edge $e \in \mathcal{E}$ by

$$\langle a, b\rangle_e := \int_e a\,b \quad \text{if } a, b \in \mathrm{T}(e), \qquad \langle \boldsymbol{a}, \boldsymbol{b}\rangle_e := \int_e \boldsymbol{a}\cdot\boldsymbol{b} \quad \text{if } \boldsymbol{a}, \boldsymbol{b} \in [\mathrm{T}(e)]^2,$$

and similarly for volume terms on an element $K \in \mathcal{T}_h$,

$$(a, b)_K := \int_K a\,b \quad \text{if } a, b \in \mathrm{H}^l(K), \qquad (\boldsymbol{a}, \boldsymbol{b})_K := \int_K \boldsymbol{a}\cdot\boldsymbol{b} \quad \text{if } \boldsymbol{a}, \boldsymbol{b} \in [\mathrm{H}^l(K)]^2.$$

If Γ is a subset of $\mathcal{E}$, we denote the L^2-norm of $q \in \mathrm{T}(\mathcal{E})$ along Γ by $\|q\|_\Gamma^2 := \sum_{e\in\Gamma} \|q\|_e^2$ and $\|q\|_e^2 := \langle q, q\rangle_e$. Similarly, if $\mathcal{T}_i$ is a subset of $\mathcal{T}_h$, we denote the L^2-norm of a $v \in \mathrm{H}^l(\mathcal{T}_i)$ by $\|v\|_{\mathcal{T}_i}^2 := \sum_{K\in\mathcal{T}_i} \|v\|_K^2$.

For $v \in \mathrm{H}^1(\mathcal{T}_h)$ we define functions whose restrictions to each element, $K \in \mathcal{T}_h$, are equal to the gradient of v. This operator in the literature is called the piecewise gradient and is usually denoted by ∇_h. For the sake of simplicity we use ∇v instead of $\nabla_h v$.

2.2. Primal formulation. To simplify our presentation, we set $\eta \ge 0$ to be a constant and $a(x) = 1$ in the model problem (1.1). Let $u, v \in \mathrm{H}^2(\mathcal{T}_h)$; then the IPH bilinear form of the model problem (1.1) is defined as

$$\begin{aligned} a(u,v) \quad := \quad & \eta\,(u,v)_{\mathcal{T}_h} + (\nabla u, \nabla v)_{\mathcal{T}_h} - \langle \{\!\{\nabla u\}\!\}, [\![v]\!]\rangle_{\mathcal{E}} - \langle \{\!\{\nabla v\}\!\}, [\![u]\!]\rangle_{\mathcal{E}} \\ & + \left\langle \tfrac{\mu}{2}[\![u]\!], [\![v]\!]\right\rangle_{\mathcal{E}} - \left\langle \tfrac{1}{2\mu}[\![\nabla u]\!], [\![\nabla v]\!]\right\rangle_{\mathcal{E}^0}, \end{aligned} \tag{2.2}$$

where $\mu \in \mathrm{T}(\mathcal{E})$, $\mu|_e = \alpha h_e^{-1}$, and $\alpha > 0$. Observe that $a(\cdot,\cdot)$ is symmetric. The definition of the IPH bilinear form is different from the classical IP method only in the last term, i.e., the last term in $a(\cdot,\cdot)$ is not present in IP.

There are two natural energy norms which are equivalent at the discrete level. Let $u \in V(h) := V_h + \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega) \subset \mathrm{H}^2(\mathcal{T}_h)$; then

$$\begin{aligned} \|u\|_{\mathtt{DG}}^2 \quad &:= \quad \eta\|u\|_{\mathcal{T}_h}^2 + \|\nabla u\|_{\mathcal{T}_h}^2 + \textstyle\sum_{e\in\mathcal{E}} \mu_e \|[\![u]\!]\|_e^2, \\ \|u\|_{\mathtt{DG},*}^2 \quad &:= \quad \|u\|_{\mathtt{DG}}^2 + \textstyle\sum_{K\in\mathcal{T}_h} h_K^2 |u|_{K,2}^2. \end{aligned} \tag{2.3}$$

One can show that they are equivalent at the discrete level by a local application of the inverse inequality (A.1).

PROPOSITION 2.1. *Let $u \in V_h$. Then we have*

$$\|u\|_{\mathtt{DG}}^2 \le \|u\|_{\mathtt{DG},*}^2 \le C^2 \|u\|_{\mathtt{DG}}^2,$$

where $C^2 > 1$ and independent of h and α.

The norm $\|\cdot\|_{\mathtt{DG},*}$ provides a natural norm for boundedness and $\|\cdot\|_{\mathtt{DG}}$ can be used for showing coercivity. The main ingredients for coercivity are the following inequalities which hold for all $u \in V_h$:

$$\begin{array}{rcl} 2\,\langle \{\!\{\nabla u\}\!\}, [\![u]\!]\rangle_{\mathcal{E}} & \le & \frac{1}{2}\|\nabla u\|_{\mathcal{T}_h}^2 + \sum_{e\in\mathcal{E}} \frac{C_1}{h_e}\|[\![u]\!]\|_e^2, \\ \left\langle \frac{1}{2\mu}[\![\nabla u]\!], [\![\nabla u]\!]\right\rangle_{\mathcal{E}^0} & \le & \frac{C_2}{\alpha}\|\nabla u\|_{\mathcal{T}_h}^2, \end{array} \tag{2.4}$$

where C_1 and C_2 are both independent of h and α but depend on the polynomial degree. This can be obtained from the trace inequality

$$\|w\|_{\partial K}^2 \le c\frac{k^2}{h}\|w\|_K^2 \quad \forall w \in \mathbb{P}^k(K), \tag{2.5}$$

where k is the polynomial degree; for details see [26, 3].

PROPOSITION 2.2. *If $\mu = \alpha h^{-1}$, for $\alpha > 0$ and sufficiently large, then we have*

$$\begin{array}{rcl} & & a(u,v) \;\le\; \overline{C}\|u\|_{\mathtt{DG},*}\|v\|_{\mathtt{DG},*} \quad \forall u,v \in V(h), \\ \underline{c}\,C^{-2}\|u\|_{\mathtt{DG},*}^2 \le \underline{c}\|u\|_{\mathtt{DG}}^2 & \le & a(u,u) \quad \forall u \in V_h, \end{array}$$

where $\underline{c} = \min\{\frac{1}{2} - \frac{C_2}{\alpha}, 1 - \frac{C_1}{\alpha}\} < 1$, $\overline{C} = 1 + \frac{C_3}{\alpha} > 1$, and both constants are independent of h.

Note that coercivity holds only for $u \in V_h$ and that $\alpha > 0$ has to be big enough to result in a positive $\underline{c}$. Since C_1 and C_2 come from the trace inequality, we can choose $\alpha = O(k^2)$, where k is the degree of the polynomials in the simplex. Throughout this paper we assume that α is chosen big enough to ensure that any term of type $1 - \frac{c}{\alpha}$ (with $c > 0$, independent of h and α) is positive.

Having established that $a(\cdot,\cdot)$ is bounded and coercive, we obtain that the following approximation problem has a unique solution: find $u_h \in V_h$ such that

$$a(u_h, v) = (f, v)_{\mathcal{T}_h} \quad \forall v \in V_h. \tag{2.6}$$

Assuming the exact solution is regular enough, it can be shown that

$$\begin{array}{rcl} \|u_h - u\|_{\mathtt{DG},*} & \le & c\,h^k |u|_{k+1,\Omega}, \\ \|u_h - u\|_0 & \le & c\,h^{k+1}|u|_{k+1,\Omega}, \end{array}$$

i.e., IPH has optimal approximation order [3, 21]. We emphasize that without setting $\mu = \alpha h^{-1}$, the coercivity and optimal approximation properties are lost.

2.3. Hybridizable formulation. In this section we exploit the fact that IPH is a hybridizable method. A method is hybridizable if one can eliminate the degrees of freedom inside each element to obtain a linear system in terms of a single-valued function along the edges, say, λ_h. Not all DG methods have this property; for example, classical IP is not hybridizable. A unified hybridization procedure for DG methods has been introduced and studied in [7], where IPH is also included.

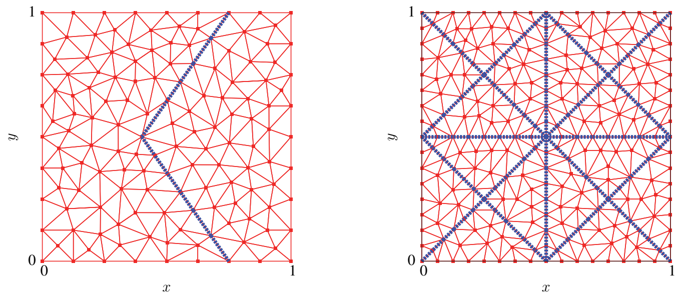

FIG. 1. *An unstructured mesh with the interface* Γ *(thick-dashed).*

We introduce the general setting by decomposing the domain into two nonoverlapping subdomains Ω_1 and Ω_2. Denoting the interface by $\Gamma := \overline{\Omega}_1 \cap \overline{\Omega}_2$, we assume $\Gamma \subset \mathcal{E}^0$, i.e., the cut does not go through any element of the triangulation. This will result in a natural partitioning of $\mathcal{T}_h$ into $\mathcal{T}_1$ and $\mathcal{T}_2$ which do not overlap but share Γ as a boundary; see Figure 1 for an example. We denote by H the maximum diameter of the subdomains and by H_Ω the diameter of the monodomain Ω. We assume $0 < h \le H < H_\Omega$.

We introduce local spaces on Ω_1 and Ω_2 by

$$V_{h,i} := \left\{ v \in \mathrm{L}^2(\Omega_i) : v|_{K \in \mathcal{T}_i} \in \mathbb{P}^k(K) \right\} \text{ for } i = 1, 2. \tag{2.7}$$

Note that this domain decomposition setting implies $V_h = V_{h,1} \oplus V_{h,2}$. We define on the interface the space of broken single-valued functions by

$$\Lambda_h := \left\{ \varphi \in \mathrm{L}^2(\Gamma) : \varphi|_{e \in \Gamma} \in \mathbb{P}^k(e) \right\}. \tag{2.8}$$

For the sake of simplicity we denote the restriction of $v \in V_h$ on $V_{h,i}$ by v_i. Observe that the trace of $v_i \in V_{h,i}$ on Γ belongs to Λ_h.

Let $(u, \lambda), (v, \varphi) \in V_h \times \Lambda_h$ and consider the symmetric bilinear form

$$\tilde{a}((u,\lambda),(v,\varphi)) := \tilde{a}_\Gamma(\lambda,\varphi) + \sum_{i=1}^{2} \left(\tilde{a}_i(u_i,v_i) + \tilde{a}_{i\Gamma}(v_i,\lambda) + \tilde{a}_{i\Gamma}(u_i,\varphi) \right), \tag{2.9}$$

where

$$\begin{aligned} \tilde{a}_\Gamma(\lambda,\varphi) &:= 2\left\langle \mu\,\lambda, \varphi \right\rangle_\Gamma, \\ \tilde{a}_{i\Gamma}(v_i,\varphi) &:= \left\langle \frac{\partial v_i}{\partial \boldsymbol{n}_i} - \mu v_i, \varphi \right\rangle_\Gamma, \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} \tilde{a}_i(u_i,v_i) \;:=\;& \eta\,(u_i,v_i)_{\mathcal{T}_i} + (\nabla u_i, \nabla v_i)_{\mathcal{T}_i} - \langle \{\!\{\nabla u_i\}\!\}, [\![v_i]\!] \rangle_{\mathcal{E}_i^0} - \langle \{\!\{\nabla v_i\}\!\}, [\![u_i]\!] \rangle_{\mathcal{E}_i^0} \\ &+ \left\langle \tfrac{\mu}{2} [\![u_i]\!], [\![v_i]\!] \right\rangle_{\mathcal{E}_i^0} - \left\langle \tfrac{1}{2\mu} [\![\nabla u_i]\!], [\![\nabla v_i]\!] \right\rangle_{\mathcal{E}_i^0} \\ &- \left\langle \frac{\partial u_i}{\partial \boldsymbol{n}_i}, v_i \right\rangle_{\partial\Omega_i} - \left\langle \frac{\partial v_i}{\partial \boldsymbol{n}_i}, u_i \right\rangle_{\partial\Omega_i} + \langle \mu\, u_i, v_i \rangle_{\partial\Omega_i}. \end{aligned} \tag{2.11}$$

This is an IPH discretization of the model problem in Ω_i and $\partial\Omega_i$ is treated as a Dirichlet boundary. Therefore $\tilde{a}_i(\cdot,\cdot)$ inherits coercivity and continuity of the original bilinear form, $a(\cdot,\cdot)$.

The global bilinear form $\tilde{a}(\cdot,\cdot)$ is also coercive at the discrete level if $\alpha > 0$ is sufficiently large, independent of h. To see this we introduce an energy norm for all $(v_i, \varphi) \in V_{h,i} \times \Lambda_h$ such that

$$\|(v_i,\varphi)\|^2_{\mathtt{HDG},i} := \eta\|v_i\|^2_{\mathcal{T}_i} + \|\nabla v_i\|^2_{\mathcal{T}_i} + \mu\|[\![v_i]\!]\|^2_{\mathcal{E}_i\setminus\Gamma} + \mu\|v_i-\varphi\|^2_{\Gamma} \quad (i=1,2). \tag{2.12}$$

Then by definition of $\tilde{a}(\cdot,\cdot)$ for all $(v,\varphi) \in V_h \times \Lambda_h$ we have

$$\begin{aligned}\tilde{a}((v,\varphi),(v,\varphi)) &= \tilde{a}_\Gamma(\varphi,\varphi) + \sum_{i=1}^2 \left(\tilde{a}_i(v_i,v_i) + 2\tilde{a}_{i\Gamma}(v_i,\varphi)\right)\\ &= \sum_{i=1}^2 \left(\tilde{a}_i(v_i,v_i) + 2\tilde{a}_{i\Gamma}(v_i,\varphi) + \frac{1}{2}\tilde{a}_\Gamma(\varphi,\varphi)\right).\end{aligned} \tag{2.13}$$

We can bound the contribution of each subdomain from below separately:

$$\begin{aligned}\tilde{a}((v,\varphi),(v,\varphi)) = \sum_{i=1}^2 &\eta\|v_i\|^2_{\mathcal{T}_i} + \|\nabla v_i\|^2_{\mathcal{T}_i}\\ &-2\left\langle \{\!\{\nabla v_i\}\!\}, [\![v_i]\!]\right\rangle_{\mathcal{E}_i\setminus\Gamma} + \frac{\mu}{2}\|[\![v_i]\!]\|^2_{\mathcal{E}_i\setminus\Gamma} - \frac{1}{2\mu}\|[\![\nabla v_i]\!]\|^2_{\mathcal{E}_i^0}\\ &-2\left\langle \frac{\partial v_i}{\partial \boldsymbol{n}_i}, v_i-\varphi\right\rangle_\Gamma + \mu\|v_i-\varphi\|^2_\Gamma\\ \geq c\sum_{i=1}^2 &\|(v_i,\varphi)\|^2_{\mathtt{HDG},i},\end{aligned}$$

where we used the inverse inequalities (2.5) for terms acting on the interface and (2.4) for terms acting inside subdomains. Here $0 < c < 1$ is a constant independent of h. Note that we proved the coercivity in a subdomain by subdomain fashion by splitting the $\tilde{a}_\Gamma(\cdot,\cdot)$ terms.

Consider the following discrete problem: find $(u_h,\lambda_h) \in V_h \times \Lambda_h$ such that

$$\tilde{a}((u_h,\lambda_h),(v,\varphi)) = (f,v)_{\mathcal{T}_h} \quad \forall (v,\varphi) \in V_h \times \Lambda_h, \tag{2.14}$$

which has a unique solution since $\tilde{a}(\cdot,\cdot)$ is coercive on $V_h \times \Lambda_h$. One can eliminate the interface variable, λ_h, and obtain a variational problem in terms of u_h only. It turns out that this coincides with the variational problem (2.6); for a proof see [21].

The advantage of the variational problem (2.14) is that each subproblem is communicating through the auxiliary unknown λ_h. Therefore we can eliminate the interior unknowns, u_i, and obtain a Schur complement system. If we test (2.14) with $v_i \neq 0$, $v_j = 0$ $(j \neq i)$, $\varphi = 0$ and assume that λ_h is known, we obtain a local problem: find $u_i \in V_{h,i}$ such that

$$\tilde{a}_i(u_i,v_i) + \tilde{a}_{i\Gamma}(v_i,\lambda_h) = (f,v_i)_{\mathcal{T}_i} \quad \forall v_i \in V_{h,i}. \tag{2.15}$$

This is an IPH discretization of the continuous problem

$$\begin{aligned}(\eta-\Delta)u &= f && \text{in } \Omega_i,\\ u &= \lambda_h && \text{on } \Gamma,\\ u &= 0 && \text{on } \partial\Omega_i \setminus \Gamma.\end{aligned}$$

However, the boundary condition on Γ is imposed weakly and therefore $u_i|_\Gamma \neq \lambda_h$ in the strong sense; see [7, 15, 21].

2.4. Schur complement formulation. We choose nodal basis functions for $\mathbb{P}^k(K)$ and denote the space of DOFs of V_h by V and similarly for subspaces by $\{V_i\}$. The variational form in (2.6) is equivalent to the linear system $A\boldsymbol{u} = \boldsymbol{f}$. A is the system matrix and $\boldsymbol{u} \in V$ are the corresponding DOFs of the approximation $u_h \in V_h$. We can partition $\boldsymbol{u}$ into $\{\boldsymbol{u}_i\}$, where $\boldsymbol{u}_i$ corresponds to DOFs of $u_i \in V_{h,i}$. Then we can arrange the entries of A and rewrite the linear system as

$$\begin{bmatrix} A_1 & A_{12} \\ A_{21} & A_2 \end{bmatrix} \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{pmatrix}. \tag{2.16}$$

We use nodal basis functions for Λ_h and denote by $\boldsymbol{\lambda}$ the corresponding DOFs for $\lambda_h \in \Lambda_h$. Then the variational form (2.14) can be written as

$$\begin{bmatrix} \tilde{A}_1 & & \tilde{A}_{1\Gamma} \\ & \tilde{A}_2 & \tilde{A}_{2\Gamma} \\ \tilde{A}_{\Gamma 1} & \tilde{A}_{\Gamma 2} & \tilde{A}_\Gamma \end{bmatrix} \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ 0 \end{pmatrix}, \tag{2.17}$$

where $\tilde{A}_{\Gamma i} = \tilde{A}_{i\Gamma}^\top$. Since this matrix is s.p.d. and the same holds also for its diagonal blocks, we can form a Schur complement system. We define $\tilde{B}_i := \tilde{A}_{\Gamma i}\tilde{A}_i^{-1}\tilde{A}_{i\Gamma}$ and $\boldsymbol{g}_\Gamma := -\sum_{i=1}^2 \tilde{A}_{\Gamma i}\tilde{A}_i^{-1}\boldsymbol{f}_i$. Then the Schur complement system reads

$$\tilde{S}_\Gamma \boldsymbol{\lambda} := \left(\tilde{A}_\Gamma - \sum_{i=1}^2 \tilde{B}_i \right) \boldsymbol{\lambda} = \boldsymbol{g}_\Gamma. \tag{2.18}$$

Definition 2.3 (discrete harmonic extension). *For all $\varphi \in \Lambda_h$, we denote by $\mathcal{H}_i(\varphi) \in V_{h,i}$ the discrete harmonic extension into Ω_i,*

$$\mathcal{H}_i(\varphi) \equiv -\tilde{A}_i^{-1}\tilde{A}_{i\Gamma}\boldsymbol{\varphi}. \tag{2.19}$$

The corresponding φ is called generator. *In other words $u_i := \mathcal{H}_i(\varphi)$ is an approximation obtained from the IPH discretization in Ω_i using φ as Dirichlet data; i.e., $\tilde{A}_i\boldsymbol{u}_i + \tilde{A}_{i\Gamma}\boldsymbol{\varphi} = 0$.*

The following result shows that an application of $\tilde{B}_i\boldsymbol{\lambda}$ can be viewed as finding the harmonic extension, $u_i := \mathcal{H}_i(\lambda_h)$, and then evaluating a "Robin-like trace" on the interface.

Proposition 2.4. *Let $\lambda_h \in \Lambda_h$ and define its harmonic extension by $u_i := \mathcal{H}_i(\lambda_h)$. Then $\boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\lambda} = \langle \mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}, \varphi\rangle_\Gamma$ for all $\varphi \in \Lambda_h$.*

Proof. Let $u_i := \mathcal{H}_i(\lambda_h)$. Then by definition of $\tilde{B}_i$ and $\tilde{a}_{i\Gamma}(\cdot,\cdot)$ we have

$$\boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\lambda} = \boldsymbol{\varphi}^\top \tilde{A}_{\Gamma i}\tilde{A}_i^{-1}\tilde{A}_{i\Gamma}\boldsymbol{\lambda} = -\boldsymbol{\varphi}^\top \tilde{A}_{\Gamma i}\boldsymbol{u}_i = \left\langle \mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}, \varphi \right\rangle_\Gamma$$

for all $\varphi \in \Lambda_h$, which completes the proof, since $\tilde{A}_{\Gamma i} = \tilde{A}_{i\Gamma}^\top$. ∎

3. Properties of the Schur complement and technical tools. The main goal of this section is to provide estimates for the minimum and maximum eigenvalues of the $\tilde{S}_\Gamma$ and $\tilde{B}_i$ for $i = 1, 2$. We use the estimate for the $\tilde{B}_i$ operators to prove convergence of the Schwarz method and provide the contraction factor later in section 4. In particular we prove in this section that the following estimates hold for all $\varphi \in \Lambda_h$:

$$c_B\, \mu\|\varphi\|_\Gamma^2 \le \boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\varphi} \le \left(1 - C_B\frac{h}{H\alpha}\right)\mu\|\varphi\|_\Gamma^2, \tag{3.1}$$

$$c\frac{H}{H_\Omega^2}\|\varphi\|_\Gamma^2 \le \boldsymbol{\varphi}^\top \tilde{S}_\Gamma \boldsymbol{\varphi} \le C\frac{\alpha}{h}\|\varphi\|_\Gamma^2, \tag{3.2}$$

where all constants are positive and independent of h, H, and H_Ω. Since $\tilde{S}_\Gamma$ and $\tilde{B}_i$ are symmetric, we can use Rayleigh quotient arguments and obtain an estimate for the minimum and maximum eigenvalues. One can also obtain an estimate with polynomial degree dependency using the techniques of this section.

The only constraint on the shape of the subdomains is a star-shape assumption. To prove the above estimates we need trace and Poincaré inequalities for totally discontinuous functions. The following trace estimate is due to Feng and Karakashian [12, Lemma 3.1]. The Poincaré inequality is due to Brenner [5].

LEMMA 3.1 (trace inequality). *Let D be a star-shape domain with diameter H_D, and triangulation $\mathcal{T}_h$. Then, for any $u \in \mathrm{H}^1(\mathcal{T}_h)$, we have*

$$\|u\|_{\partial D}^2 \le c\left[H_D^{-1}\|u\|_D^2 + H_D\left(\|\nabla u\|_D^2 + h^{-1}\|[\![u]\!]\|_{\mathcal{E}\setminus\partial D}^2\right)\right].$$

LEMMA 3.2 (Poincaré inequality). *Let D be an open connected polygonal domain with diameter H_D and triangulation $\mathcal{T}_h$. Then, for any $u \in \mathrm{H}^1(\mathcal{T}_h)$ we have*

$$\|u\|_D^2 \le cH_D^2\left[\|\nabla u\|_D^2 + h^{-1}\|[\![u]\!]\|_{\mathcal{E}\setminus\partial D}^2 + h^{-1}\|u\|_\nu^2\right],$$

where ν is a measurable subset of ∂D with nonzero measure.

3.1. Eigenvalue estimates for $\tilde{B}_i$. In order to obtain estimates for the eigenvalues of the $\tilde{B}_i$ operator, we first recall Definition 2.3 of a harmonic extension: $u_i \in V_{h,i}$ is called harmonic extension of $\varphi \in \Lambda_h$ if it satisfies $\tilde{A}_i\boldsymbol{u}_i + \tilde{A}_{i\Gamma}\boldsymbol{\varphi} = 0$. Now multiplying this relation by $\boldsymbol{u}_i^\top$ from the left we get

$$\begin{array}{rll} & \boldsymbol{u}_i^\top\tilde{A}_i\boldsymbol{u}_i + \boldsymbol{u}_i^\top\tilde{A}_{i\Gamma}\boldsymbol{\varphi} & = \quad 0 \\ \Leftrightarrow & \boldsymbol{u}_i^\top\tilde{A}_i\boldsymbol{u}_i - \boldsymbol{\varphi}^\top\tilde{A}_{\Gamma i}\tilde{A}_i^{-1}\tilde{A}_{i\Gamma}\boldsymbol{\varphi} & = \quad 0 \\ \Leftrightarrow & \boldsymbol{u}_i^\top\tilde{A}_i\boldsymbol{u}_i - \boldsymbol{\varphi}^\top\tilde{B}_i\boldsymbol{\varphi} & = \quad 0, \end{array}$$

where we used $\boldsymbol{u}_i = -\tilde{A}_i^{-1}\tilde{A}_{i\Gamma}\boldsymbol{\varphi}$, $\tilde{A}_{\Gamma i} = \tilde{A}_{i\Gamma}^\top$, and the definition of $\tilde{B}_i$. Hence if $u_i = \mathcal{H}_i(\varphi)$, then we have

$$\boldsymbol{\varphi}^\top\tilde{B}_i\boldsymbol{\varphi} = \tilde{a}_i(u_i, u_i). \tag{3.3}$$

Now recall that $\tilde{a}_i(\cdot,\cdot)$ is coercive and bounded over $V_{h,i}$; therefore $\underline{c}\|u_i\|_{\mathtt{DG}}^2 \le \tilde{a}_i(u_i,u_i) \le \overline{C}\|u_i\|_{\mathtt{DG}}^2$. Thus if we relate the energy norm of the harmonic extension, $u_i := \mathcal{H}_i(\varphi) \in V_{h,i}$, to the L^2-norm of φ we obtain the desired estimate (3.1). More precisely we can show that the estimate

$$c_{\mathcal{H}} \cdot \mu\|\varphi\|_\Gamma^2 \le \|u_i\|_{\mathtt{DG}}^2 \le C_{\mathcal{H}} \cdot \mu\|\varphi\|_\Gamma^2 \tag{3.4}$$

holds, where $0 < c_{\mathcal{H}} < 1$ and $C_{\mathcal{H}} > 1$ are constants independent of h. Observe that $C_{\mathcal{H}} > 1$, while the upper bound estimate in (3.1) is less than one. We show later how one can obtain a sharp upper bound estimate as in (3.1).

Let us start with the lower bound of inequality (3.4). First we introduce an extension by zero operator $\boldsymbol{\theta}_i : \Lambda_h \to V_{h,i}$ which is defined for all $\varphi \in \Lambda_h$ as

$$\boldsymbol{\theta}_i(\varphi) := \begin{cases} \varphi & \text{on edges belonging to } \Gamma, \\ 0 & \text{on other nodes.} \end{cases}$$

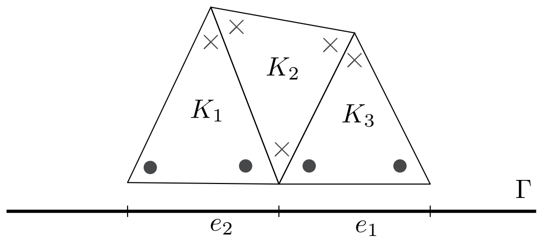

FIG. 2. *Illustration of the extension by zero, $\boldsymbol{\theta}_i(\varphi)$, for elements which share an edge with the interface, e.g., $\{K_1, K_3\}$, and those which do not, e.g., K_2.*

For a graphical illustration see Figure 2. Note that there are elements like K_2 which physically share a node and not an *edge* with the interface, but we leave $\boldsymbol{\theta}_i(\varphi)$ in K_2 to be zero. More precisely, only those elements which share an edge with the interface are nonzero.

We show in the appendix (see also [23]) that in an element, $K \in \mathcal{T}_i$, with an edge $e \in \Gamma$ we have

$$\begin{aligned} \|\boldsymbol{\theta}_i(\varphi)\|_K^2 &\le C_3\, h\|\varphi\|_e^2, \\ \|\nabla\boldsymbol{\theta}_i(\varphi)\|_K^2 &\le C_4 h^{-1}\|\varphi\|_e^2, \\ \|[\![\boldsymbol{\theta}_i(\varphi)]\!]\|_{\mathcal{E}_i}^2 &\le C_5\|\varphi\|_\Gamma^2, \end{aligned} \tag{3.5}$$

where $C_3 > 0$, $C_4 > 0$, and $C_5 \ge 1$ and all are independent of h. This yields the following result, which relates the energy of the extension by zero to its L^2-norm on the interface.

LEMMA 3.3. *Let $\varphi \in \Lambda_h$ and $\boldsymbol{\theta}_i(\varphi)$ be its extension by zero into Ω_i. We have*

$$\|\boldsymbol{\theta}_i(\varphi)\|_{\mathtt{DG}}^2 \le \mu\, C_\theta\, \|\varphi\|_\Gamma^2,$$

where $C_\theta = C_3\eta + C_4\alpha^{-1} + C_5 > 1$.

Proof. First note that by definition $\boldsymbol{\theta}_i(\varphi)$ and $\nabla\boldsymbol{\theta}_i(\varphi)$ are nonzero only on those elements which share an edge with the interface. We call them $\{K_\Gamma\} \subset \mathcal{T}_i$. Then we have

$$\begin{aligned} \|\boldsymbol{\theta}_i(\varphi)\|_{\mathtt{DG}}^2 &= \sum_{K\in\{K_\Gamma\}} \eta\|\boldsymbol{\theta}_i(\varphi)\|_K^2 + \|\nabla\boldsymbol{\theta}_i(\varphi)\|_K^2 + \mu\|[\![\boldsymbol{\theta}_i(\varphi)]\!]\|_{\mathcal{E}_i}^2 \\ &\le C_3\,\eta\, h\|\varphi\|_\Gamma^2 + \frac{C_4}{h}\|\varphi\|_\Gamma^2 + C_5\,\mu\|\varphi\|_\Gamma^2 \\ &\le \mu\left(C_3\,\eta + \frac{C_4}{\alpha} + C_5\right)\|\varphi\|_\Gamma^2, \end{aligned}$$

which completes the proof with $C_\theta := C_3\,\eta + \frac{C_4}{\alpha} + C_5 > 1$. ∎

Now we are able to relate the energy of a harmonic extension, $u_i := \mathcal{H}_i(\varphi)$, to the L^2-norm of φ on the interface.

LEMMA 3.4. *Let $\varphi \in \Lambda_h$ and $u_i := \mathcal{H}_i(\varphi)$ be its harmonic extension into Ω_i. Then we have*

$$c_{\mathcal{H}} \cdot \mu\|\varphi\|_\Gamma^2 \le \|u_i\|_{\mathtt{DG}}^2,$$

where $c_{\mathcal{H}} = (1 - \frac{c}{\alpha})^2 \cdot \frac{1}{C_\theta \overline{C}^2} < 1$.

Proof. Since u_i is the harmonic extension of φ, it satisfies (2.15) (with $f = 0$). Let $v = \boldsymbol{\theta}_i(\varphi)$. Then by definition of $\tilde{a}_{i\Gamma}(\cdot,\cdot)$ we have

$$\tilde{a}_i(u_i, \boldsymbol{\theta}_i(\varphi)) = \left\langle \mu\, \boldsymbol{\theta}_i(\varphi) - \frac{\partial \boldsymbol{\theta}_i(\varphi)}{\partial \boldsymbol{n}_i}, \varphi \right\rangle_\Gamma .$$

Note that $\boldsymbol{\theta}_i(\varphi)|_\Gamma = \varphi$. We can bound the right-hand side from below; therefore

$$\begin{aligned}
\tilde{a}_i(u_i, \boldsymbol{\theta}_i(\varphi)) &\geq \mu\|\varphi\|_\Gamma^2 - \left\| \frac{\partial \boldsymbol{\theta}_i(\varphi)}{\partial \boldsymbol{n}_i} \right\|_\Gamma \|\varphi\|_\Gamma \\
&\geq \mu\|\varphi\|_\Gamma^2 - \frac{c}{\sqrt{h}} \|\nabla \boldsymbol{\theta}_i(\varphi)\|_{K_\Gamma} \|\varphi\|_\Gamma \quad \text{by inequality (2.5)} \\
&\geq \mu\|\varphi\|_\Gamma^2 - \frac{c'}{h} \|\varphi\|_\Gamma^2 \quad \text{by inequality (3.5)} \\
&= \mu \left(1 - \frac{c'}{\alpha}\right) \|\varphi\|_\Gamma^2,
\end{aligned}$$

which is positive if $\alpha > 0$ and sufficiently large. By continuity of $\tilde{a}_i(\cdot,\cdot)$ we have

$$\mu \left(1 - \frac{c'}{\alpha}\right) \|\varphi\|_\Gamma^2 \leq \overline{C}\, \|u_i\|_{\mathtt{DG}} \cdot \|\boldsymbol{\theta}_i(\varphi)\|_{\mathtt{DG}}.$$

Note that we are able to use $\|\cdot\|_{\mathtt{DG}}$ instead of $\|\cdot\|_{\mathtt{DG},*}$ since we work with discrete spaces. An application of Lemma 3.3 completes the proof with $c_{\mathcal{H}} = (1 - \frac{c'}{\alpha})^2 \cdot \frac{1}{C_\theta \overline{C}^2} < 1$. ☐

The upper bound in (3.4) can be obtained much easier using coercivity of the $\tilde{a}_i(\cdot,\cdot)$.

LEMMA 3.5. *Let $\varphi \in \Lambda_h$ and $u_i := \mathcal{H}_i(\varphi)$ be its harmonic extension into Ω_i. Then we have*

$$\|u_i\|_{\mathtt{DG}}^2 \leq C_{\mathcal{H}} \cdot \mu\|\varphi\|_\Gamma^2,$$

where $C_{\mathcal{H}} = (1 + \frac{C}{\sqrt{\alpha}})^2 \cdot \frac{1}{\underline{c}^2} > 1$.

Proof. Since u_i is the harmonic extension of φ, it satisfies (2.15) (with $f = 0$). Using the fact that $\tilde{a}_i(\cdot,\cdot)$ is coercive we have

$$\begin{aligned}
\underline{c}\|u_i\|_{\mathtt{DG}}^2 \leq \tilde{a}_i(u_i, u_i) &= -\tilde{a}_{i\Gamma}(u_i, \varphi) \\
&= \left\langle \mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}, \varphi \right\rangle_\Gamma \\
&\leq \mu\|u_i\|_\Gamma \|\varphi\|_\Gamma + \left\| \frac{\partial u_i}{\partial \boldsymbol{n}_i} \right\|_\Gamma \|\varphi\|_\Gamma \\
&\leq \mu\|u_i\|_\Gamma \|\varphi\|_\Gamma + \frac{C}{\sqrt{h}} \|\nabla u_i\|_{\mathcal{T}_i} \|\varphi\|_\Gamma \\
&\leq \mu^{\frac{1}{2}} \left(1 + \frac{C}{\sqrt{\alpha}}\right) \|u_i\|_{\mathtt{DG}} \cdot \|\varphi\|_\Gamma,
\end{aligned}$$

which completes the proof with $C_{\mathcal{H}} := (1 + \frac{C}{\sqrt{\alpha}})^2 \cdot \frac{1}{\underline{c}^2} > 1$. ☐

We see that $C_{\mathcal{H}} > 1$, which does not provide a sharp estimate for the maximum eigenvalue of $\tilde{B}_i$. We now show how to obtain a sharp estimate for the maximum eigenvalue of the $\tilde{B}_i$. Recall that the global matrix $\tilde{A}$ is s.p.d. and the positive definiteness is proved by using for each subdomain $\frac{1}{2}\tilde{A}_\Gamma$ in (2.13). Therefore we consider the s.p.d. matrix

$$\hat{A} := \begin{bmatrix} \tilde{A}_i & \tilde{A}_{i\Gamma} \\ \tilde{A}_{\Gamma i} & \frac{1}{2}\tilde{A}_\Gamma \end{bmatrix}.$$

To show positive-definiteness, let $\boldsymbol{w} := (\boldsymbol{u}_i, \boldsymbol{\varphi})^\top$ and observe

$$\boldsymbol{w}^\top \hat{A}\boldsymbol{w} = \tilde{a}_i(u_i, u_i) + 2\tilde{a}_{i\Gamma}(u_i, \varphi) + \frac{1}{2}\tilde{a}_\Gamma(\varphi, \varphi) \geq c\|(u_i, \varphi)\|^2_{\mathtt{HDG},i} \tag{3.6}$$

for all $u_i \in V_{h,i}$ and $\varphi \in \Lambda_h$. Now let $u_i = \mathcal{H}_i(\varphi)$; then by a simple manipulation we have $\boldsymbol{\varphi}^\top(\frac{1}{2}\tilde{A}_\Gamma - \tilde{B}_i)\boldsymbol{\varphi} = \boldsymbol{w}^\top \hat{A}\boldsymbol{w}$. Combining with (3.6) and recalling that $\boldsymbol{\varphi}^\top \tilde{A}_\Gamma \boldsymbol{\varphi} = 2\mu\|\varphi\|^2_\Gamma$ we obtain

$$\mu\|\varphi\|^2_\Gamma - c\|(\mathcal{H}_i(\varphi), \varphi)\|^2_{\mathtt{HDG},i} \geq \boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\varphi}. \tag{3.7}$$

This gives a sharp estimate for the maximum eigenvalue of $\tilde{B}_i$ if we can bound the second term from below, which is stated in the following lemma.

LEMMA 3.6. *Let $\varphi \in \Lambda_h$ and $u_i \in V_{h,i}$ for $i = 1, 2$. Let H_i be the diameter of the subdomain. Then we have*

$$\frac{c}{H_i}\|\varphi\|^2_\Gamma \leq \|(u_i, \varphi)\|^2_{\mathtt{HDG},i}.$$

Proof. We first invoke, triangle inequality and then Young's inequality,

$$\|\varphi\|^2_\Gamma \leq \|u_i - \varphi\|^2_\Gamma + \|u_i\|^2_\Gamma \leq H_i h^{-1}\|u_i - \varphi\|^2_\Gamma + \|u_i\|^2_\Gamma,$$

where the last inequality is due to the fact that $h \leq H_i$. Now for the second term on the right-hand side we apply the trace inequality from Lemma 3.1 and subsequently the Poincaré inequality from Lemma 3.2 with $\nu = \partial\Omega_i \setminus \Gamma$. We obtain

$$\begin{aligned} \|\varphi\|^2_\Gamma &\leq H_i h^{-1}\|u_i - \varphi\|^2_\Gamma + c_1 H_i \left(\|\nabla u_i\|^2_{\Omega_i} + h^{-1}\|[\![u_i]\!]\|^2_{\mathcal{E}_i \setminus \Gamma}\right) \\ &\leq c_2 H_i \|(u_i, \varphi)\|^2_{\mathtt{HDG},i}, \end{aligned}$$

which completes the proof. ◻

We are now in position to prove the estimate for the eigenvalues of $\tilde{B}_i$.

LEMMA 3.7. *There exists $\alpha > 0$, sufficiently large, such that*

$$c_B\mu\|\varphi\|^2_\Gamma \leq \boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\varphi} \leq \left(1 - C_B\frac{h}{H\alpha}\right)\mu\|\varphi\|^2_\Gamma \quad \forall \varphi \in \Lambda_h,$$

where $0 < c_B < 1$. Therefore $\tilde{B}_i$ is s.p.d. Moreover $\tilde{A}_\Gamma - 2\tilde{B}_i$ is s.p.d.

Proof. To show the proof of the lower bound we use (3.3), coercivity of $\tilde{a}_i(\cdot, \cdot)$ and Lemma 3.4 to obtain

$$\boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\varphi} = \tilde{a}_i(u_i, u_i) \geq \underline{c}\|u_i\|^2_{\mathtt{DG}} \geq \underline{c} \cdot c_{\mathcal{H}} \cdot \mu\|\varphi\|^2_\Gamma.$$

This completes the lower bound by setting $c_B := \underline{c} \cdot c_{\mathcal{H}} < 1$. For the upper bound we use inequality (3.7) and Lemma 3.6, where we obtain

$$\boldsymbol{\varphi}^\top \tilde{B}_i \boldsymbol{\varphi} \leq \left(1 - \frac{c}{H}\frac{h}{\alpha}\right)\mu\|\varphi\|^2_\Gamma.$$

Finally from inequality (3.7) we have that $\tilde{A}_\Gamma - 2\tilde{B}_i$ is s.p.d. ∎

Remark 3.8. This estimate shows that the condition number satisfies

$$\kappa(\tilde{B}_i) \le c_B^{-1}\left(1 - C_B\frac{h}{H\alpha}\right),$$

which implies that $\tilde{B}_i$ is scalable. In other words, if we keep the ratio h/H constant the condition number does not change. Geometrically that is equivalent to scaling the subdomain and the triangulation at the same rate which does not change the entries of the $\tilde{B}_i$ nor its size. Therefore the condition number of $\tilde{B}_i$ is expected not to change.

3.2. Eigenvalue estimate for $\tilde{S}_\Gamma$. Estimating eigenvalues of the Schur complement is similar to estimating eigenvalues of $\tilde{B}_i$. To show the lower bound in estimate (3.2), we need the following lemma.

LEMMA 3.9. *Let $\varphi \in \Lambda_h$ and $u_i \in V_{h,i}$ for $i = 1, 2$. Let H_Ω be the diameter of the domain and H be the maximum diameter of the subdomains. Then we have*

$$c\frac{H}{H_\Omega^2}\|\varphi\|_\Gamma^2 \le \sum_{i=1}^{2}\|(u_i,\varphi)\|_{\mathtt{HDG},i}^2.$$

Proof. First we invoke a triangle inequality

$$H_i\|\varphi\|_\Gamma^2 \le H_i\|u_i - \varphi\|_\Gamma^2 + H_i\|u_i\|_\Gamma^2 \le H_i^2 h^{-1}\|u_i - \varphi\|_\Gamma^2 + H_i\|u_i\|_\Gamma^2,$$

where the last inequality is due to the fact that $h \le H_i$. Now for the second term on the right-hand side, observe that using Lemma 3.1 we have

$$c_i H_i\|u_i\|_\Gamma^2 \le c_i H_i\|u_i\|_{\partial\Omega_i}^2 \le \|u_i\|_{\Omega_i}^2 + H_i^2\left(\|\nabla u_i\|_{\Omega_i}^2 + h^{-1}\|[\![u_i]\!]\|_{\mathcal{E}_i\setminus\partial\Omega_i}^2\right).$$

We sum over both subdomains and invoke Lemma 3.2 for the L^2-norm of u over Ω

$$\begin{aligned} cH\sum_{i=1}^{2}\|u_i\|_\Gamma^2 &\le \|u\|_\Omega^2 + H^2\sum_{i=1}^{2}\left(\|\nabla u_i\|_{\Omega_i}^2 + h^{-1}\|[\![u_i]\!]\|_{\mathcal{E}_i\setminus\partial\Omega_i}^2\right) \\ &\le CH_\Omega^2\left(\|\nabla u\|_\Omega^2 + h^{-1}\|[\![u]\!]\|_{\mathcal{E}\setminus\partial\Omega}^2 + h^{-1}\|u\|_{\partial\Omega}^2\right) \\ &\quad + H^2\sum_{i=1}^{2}\left(\|\nabla u_i\|_{\Omega_i}^2 + h^{-1}\|[\![u_i]\!]\|_{\mathcal{E}_i\setminus\partial\Omega_i}^2\right). \end{aligned}$$

Noting that $H \le H_\Omega$ and by definition of $\|(u_i,\varphi)\|_{\mathtt{HDG},i}$ we obtain

$$\begin{aligned} cH\sum_{i=1}^{2}\|u_i\|_\Gamma^2 &\le H_\Omega^2\sum_{i=1}^{2}\left(\|\nabla u_i\|_{\Omega_i}^2 + h^{-1}\|[\![u_i]\!]\|_{\mathcal{E}_i\setminus\partial\Omega_i}^2 + h^{-1}\|u_i\|_{\partial\Omega\cap\partial\Omega_i}^2\right) \\ &\quad + H_\Omega^2 h^{-1}\|[\![u]\!]\|_\Gamma^2 \\ &\le H_\Omega^2\sum_{i=1}^{2}\left(\|\nabla u_i\|_{\Omega_i}^2 + h^{-1}\|[\![u_i]\!]\|_{\mathcal{E}_i\setminus\partial\Omega_i}^2 + h^{-1}\|u_i\|_{\partial\Omega\cap\partial\Omega_i}^2\right) \\ &\quad + H_\Omega^2 h^{-1}\left(\|u_1 - \varphi\|_\Gamma^2 + \|u_2 - \varphi\|_\Gamma^2\right) \\ &\le H_\Omega^2\sum_{i=1}^{2}\|(u_i,\varphi)\|_{\mathtt{HDG},i}^2. \end{aligned}$$

Substituting back into the first inequality completes the proof. ∎

LEMMA 3.10. *There exists $\alpha > 0$, sufficiently large, such that*

$$c\frac{H}{H_\Omega^2}\|\varphi\|_\Gamma^2 \le \boldsymbol{\varphi}^\top\tilde{S}_\Gamma\boldsymbol{\varphi} \le \frac{2\alpha}{h}\|\varphi\|_\Gamma^2.$$

Therefore $\tilde{S}_\Gamma$ is s.p.d. Moreover $\tilde{A}_\Gamma - \tilde{B}_i$ is s.p.d.

Proof. The symmetry is easy to check since $\tilde{A}_\Gamma$ and $\tilde{B}_1$, $\tilde{B}_2$ are symmetric. For the upper bound in the estimate we recall that $\tilde{B}_1$, $\tilde{B}_2$ are positive definite and hence

$$\boldsymbol{\varphi}^\top \tilde{S}_\Gamma \boldsymbol{\varphi} = \boldsymbol{\varphi}^\top \left(\tilde{A}_\Gamma - \sum_{i=1}^{2} \tilde{B}_i \right) \boldsymbol{\varphi} \le \boldsymbol{\varphi}^\top \tilde{A}_\Gamma \boldsymbol{\varphi} = 2\mu \|\varphi\|_\Gamma^2.$$

Now let $u_i := \mathcal{H}_i(\varphi)$ and $\boldsymbol{v} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{\varphi})^\top$. A straightforward calculation shows that $\boldsymbol{\varphi}^\top \tilde{S}_\Gamma \boldsymbol{\varphi} = \boldsymbol{v}^\top \tilde{A} \boldsymbol{v}$. Then the coercivity of the bilinear form $\tilde{a}(\cdot,\cdot)$ and an application of Lemma 3.9 yields

$$\boldsymbol{\varphi}^\top \tilde{S}_\Gamma \boldsymbol{\varphi} = \boldsymbol{v}^\top \tilde{A} \boldsymbol{v} \equiv \tilde{a}((u,\varphi),(u,\varphi)) \ge c \sum_{i=1}^{2} \|(u_i,\varphi)\|_{\mathtt{HDG},i}^2 \ge c \frac{H}{H_\Omega^2} \|\varphi\|_\Gamma^2.$$

For the final statement, observe that for all $\varphi \neq 0$ we have

$$\boldsymbol{\varphi}^\top (\tilde{A}_\Gamma - \tilde{B}_i) \boldsymbol{\varphi} > \boldsymbol{\varphi}^\top \left(\tilde{A}_\Gamma - \sum_{j=1}^{2} \tilde{B}_j \right) \boldsymbol{\varphi} > 0,$$

since the $\{\tilde{B}_i\}$ are positive definite. This completes the proof. ∎

Remark 3.11. Note that Lemma 3.10 provides an upper bound for the condition number, $\kappa(\tilde{S}_\Gamma) \le O(\frac{\alpha}{h})$. A similar result also holds for the classical FEM; see [4] and [25, Lemma 4.11].

4. Schwarz methods and the Schur complement.

In order to solve the Schur complement system we can devise a Schwarz method to obtain λ_h. We will prove that a natural Schwarz method for the Schur complement is equivalent to the block Jacobi iteration in (1.3), but it suffers from slow convergence. Later we show how to obtain an optimized Schwarz method (OSM) for the Schur complement which converges much faster to the same fixed point.

Let us relax the constraint that λ_h is single-valued. Let $\lambda_{h,1}, \lambda_{h,2} \in \Lambda_h$. Assume $\lambda_{h,2}$ is known; that is, we know $u_2 \in V_{h,2}$. Then we can split the Schur complement system (2.18) and obtain an approximation for $\lambda_{h,1}$ and consequently $u_1 \in V_{h,1}$ from

$$(\tilde{A}_\Gamma - \tilde{B}_1)\boldsymbol{\lambda}_1 = \tilde{B}_2 \boldsymbol{\lambda}_2 + \boldsymbol{g}_\Gamma.$$

As a consequence of Lemma 3.10, $(\tilde{A}_\Gamma - \tilde{B}_1)$ is invertible and we can obtain $\lambda_{h,1}$. This suggests an iterative method to obtain λ_h. We will see that this produces identical iterates as the block Jacobi method.

ALGORITHM 1. *Let $\lambda_{h,1}^{(0)}, \lambda_{h,2}^{(0)} \in \Lambda_h$ be two random initial guesses. Then for $n = 1, 2, \ldots$ find $\{\lambda_{h,i}^{(n)}\}$ such that*

$$\begin{aligned} (\tilde{A}_\Gamma - \tilde{B}_1)\boldsymbol{\lambda}_1^{(n)} &= \tilde{B}_2 \boldsymbol{\lambda}_2^{(n-1)} + \boldsymbol{g}_\Gamma, \\ (\tilde{A}_\Gamma - \tilde{B}_2)\boldsymbol{\lambda}_2^{(n)} &= \tilde{B}_1 \boldsymbol{\lambda}_1^{(n-1)} + \boldsymbol{g}_\Gamma. \end{aligned} \tag{4.1}$$

At convergence, we have $\tilde{A}_\Gamma(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2) = 0$, which implies $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \tilde{S}_\Gamma^{-1} \boldsymbol{g}_\Gamma$.

The following result shows that the above method generates the same iterates as the block Jacobi iteration (1.3). By linearity it suffices to consider the error equation, $f = 0$, which implies $\boldsymbol{g}_\Gamma = 0$.

Proposition 4.1. *Let $\lambda_{h,1}^{(0)}, \lambda_{h,2}^{(0)}$ be two random initial guesses of Algorithm* 1 *and without loss of generality suppose $f = 0$. Set the initial guess of the block Jacobi iteration* (1.3) *to be $u_i^{(0)} = \mathcal{H}_i(\lambda_{h,i}^{(0)})$. Then $u_i^{(n)} = \mathcal{H}_i(\lambda_{h,i}^{(n)})$ for all $n > 0$, i.e., both methods produce the same iterates.*

Proof. See [20] for the proof. ∎

4.1. Analysis of classical Schwarz for the Schur complement. By linearity we consider the error equations and we denote by $\boldsymbol{e}_i^{(n)} := \boldsymbol{\lambda}_i^{(n)} - \boldsymbol{\lambda}$. The iterations in (4.1) can be rewritten in a more suitable form for analysis. Since $\tilde{A}_\Gamma$ is s.p.d. (it is just a scaled mass matrix), the square root $\tilde{A}_\Gamma^{1/2}$ exists and is also s.p.d. Therefore, for $i, j \in \{1, 2\}$ and $i \neq j$ we can write equivalently

$$\begin{aligned} (\tilde{A}_\Gamma - \tilde{B}_i)\boldsymbol{e}_i^{(n)} &= \tilde{B}_j \boldsymbol{e}_j^{(n-1)} \\ \Leftrightarrow \tilde{A}_\Gamma^{1/2}(I - \tilde{A}_\Gamma^{-1/2}\tilde{B}_i\tilde{A}_\Gamma^{-1/2})\tilde{A}_\Gamma^{1/2}\boldsymbol{e}_i^{(n)} &= \tilde{B}_j \boldsymbol{e}_j^{(n-1)} \\ \Leftrightarrow \qquad (I - \tilde{A}_\Gamma^{-1/2}\tilde{B}_i\tilde{A}_\Gamma^{-1/2})\tilde{\boldsymbol{e}}_i^{(n)} &= (\tilde{A}_\Gamma^{-1/2}\tilde{B}_j\tilde{A}_\Gamma^{-1/2})\tilde{\boldsymbol{e}}_j^{(n-1)}, \end{aligned}$$

where $\tilde{\boldsymbol{e}}_i = \tilde{A}_\Gamma^{1/2}\boldsymbol{e}_i$. We define

$$C_i := \tilde{A}_\Gamma^{-1/2}\tilde{B}_i\tilde{A}_\Gamma^{-1/2}, \tag{4.2}$$

which is invertible and symmetric. Since $\tilde{A}_\Gamma - \tilde{B}_i$ is invertible and $\tilde{A}_\Gamma^{1/2}$ exists we can conclude that $I - C_i$ is also invertible by definition. Therefore we have

$$(I - C_i)\tilde{\boldsymbol{e}}_i^{(n)} = C_j\tilde{\boldsymbol{e}}_j^{(n-1)} = C_j(I - C_j)^{-1}C_i\tilde{\boldsymbol{e}}_i^{(n-2)}$$

or

$$\boldsymbol{\varphi}_i^{(n)} = C_j(I - C_j)^{-1} \cdot C_i(I - C_i)^{-1}\boldsymbol{\varphi}_i^{(n-2)},$$

where $\boldsymbol{\varphi}_i = (I - C_i)\tilde{\boldsymbol{e}}_i$. Finally the iterations can be rewritten as

$$\boldsymbol{\varphi}_i^{(n)} = (C_j^{-1} - I)^{-1} \cdot (C_i^{-1} - I)^{-1}\boldsymbol{\varphi}_i^{(n-2)}. \tag{4.3}$$

We show how the contraction factor of the iteration in (4.3) is related to the eigenvalues of $\{C_i\}$. Let $\|\cdot\|_2$ be the usual 2-norm in $\mathbb{R}^n$, and denote by $D_i := (C_i^{-1} - I)^{-1}$. Then we can estimate

$$\|\boldsymbol{\varphi}_i^{(n)}\|_2 \le \|D_jD_i\|_2\,\|\boldsymbol{\varphi}_i^{(n-2)}\|_2 \le \|D_j\|_2\,\|D_i\|_2\,\|\boldsymbol{\varphi}_i^{(n-2)}\|_2 = \rho(D_j)\,\rho(D_i)\,\|\boldsymbol{\varphi}_i^{(n-2)}\|_2,$$

since $\{D_i\}$ are symmetric. In other words we have used a different norm for the error: with $E_i := (I - C_i)\tilde{A}_\Gamma^{1/2}$, which is invertible, we have

$$\|\boldsymbol{e}_i\|_{E_i^\top E_i} = \|E_i\boldsymbol{e}_i\|_2 = \|\boldsymbol{\varphi}_i\|_2.$$

Let $\sigma(M)$ denote an eigenvalue of a given matrix M. Then we have

$$\rho(D_i) := \max_{\sigma(D_i)} |\sigma(D_i)| = \max_{\sigma(C_i)} \left|\frac{\sigma(C_i)}{1 - \sigma(C_i)}\right|.$$

Hence a sufficient condition for convergence is that $\sigma(C_i) \in (-\infty, 1/2)$. On the other hand by the definition of C_i we know that $\sigma(C_i)$ are the eigenvalues of the generalized eigenvalue problem $\tilde{B}_i\boldsymbol{\varphi} = \sigma\,\tilde{A}_\Gamma\boldsymbol{\varphi}$. Since both $\tilde{A}_\Gamma$ and $\tilde{B}_i$ are s.p.d., $\sigma(C_i)$ is positive. Therefore a sufficient condition for convergence is to show that $\sigma(C_i) \in (0, 1/2)$.

Recall that since C_i is symmetric we have

$$\sigma_{\min}(C_i) = \inf_{\boldsymbol{\varphi}\neq 0} \frac{\boldsymbol{\varphi}^\top \tilde{B}_i\boldsymbol{\varphi}}{\boldsymbol{\varphi}^\top \tilde{A}_\Gamma\boldsymbol{\varphi}} = \inf_{\boldsymbol{\varphi}\neq 0} \frac{\boldsymbol{\varphi}^\top \tilde{B}_i\boldsymbol{\varphi}}{2\mu\|\varphi\|_\Gamma^2} \geq \frac{c_B}{2}, \tag{4.4}$$

where we have used the lower bound estimate of Lemma 3.7. Here $0 < c_B < 1$. The upper bound for $\sigma_{\max}(C_i)$ can also be obtained using Lemma 3.7. Hence

$$\sigma_{\max}(C_i) = \sup_{\boldsymbol{\varphi}\neq 0} \frac{\boldsymbol{\varphi}^\top \tilde{B}_i\boldsymbol{\varphi}}{2\mu\|\varphi\|_\Gamma^2} \leq \frac{1}{2}\left(1 - C\frac{h}{\alpha H}\right), \tag{4.5}$$

which is strictly less than $\frac{1}{2}$. Consequently for the eigenvalues of D_i, we obtain the estimate

$$0 < \frac{c_B}{2 - c_B} \leq \sigma(D_i) \leq 1 - C\frac{h}{\alpha H} < 1.$$

We summarize the convergence result in the following theorem.

THEOREM 4.2. *There exists an $\alpha > 0$ independent of H and h such that Algorithm* 1 *converges and the contraction factor is bounded by*

$$\rho \leq 1 - O(h). \tag{4.6}$$

4.2. Analysis of an OSM for the Schur complement. As has been shown in [15], the IPH discretization is imposing Robin transmission conditions between subdomains, and the Robin parameter is precisely the penalty parameter μ of the DG method. For approximation purposes and ensuring coercivity, μ is set to be αh^{-1} for some $\alpha > 0$ large and independent of h.

In the Schwarz theory with Robin transmission conditions this choice of μ corresponds to damping high frequencies of the DtN operator. In other words, the low frequencies are responsible for the slow convergence of the algorithm that we have analyzed in the previous subsection, as we have shown the contraction factor is $\rho = 1 - O(h)$. Optimized Schwarz theory suggests choosing the Robin parameter $O(h^{-1/2})$ (see [13]), while this is not possible for an IPH discretization since we lose coercivity and optimal approximation properties.

The remedy comes from an idea first introduced in [8] and later independently in [10] for Maxwell's equations. The idea is to perturb the transmission conditions such that while iterating we produce a different sequence but obtain the same fixed point as the original Schwarz algorithm.

Let us introduce two new unknowns, one for each subdomain, along the interface called $\{r_{12}, r_{21}\}$ such that $r_{ij} \in \Lambda_h$. Recall that by Proposition 2.4 an application of $\tilde{B}_i\boldsymbol{\lambda}_i$ is equivalent to $\mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}$ on the interface, where $u_i := \mathcal{H}_i(\lambda_{h,i})$. Now let $r_{ij} = (\mu u_j - \frac{\partial u_j}{\partial \boldsymbol{n}_j})|_\Gamma$. Let us denote by M_Γ the mass matrix along the interface and $\boldsymbol{r}_{ij}$ the corresponding DOFs of r_{ij}. Then we observe that

$$\boldsymbol{\varphi}^\top M_\Gamma \boldsymbol{r}_{ij} = \langle r_{ij}, \varphi\rangle_\Gamma = \left\langle \mu u_j - \frac{\partial u_j}{\partial \boldsymbol{n}_j}, \varphi \right\rangle_\Gamma = \boldsymbol{\varphi}^\top \tilde{B}_j\boldsymbol{\lambda}_j \quad \forall \varphi \in \Lambda_h.$$

Therefore we conclude that

$$M_\Gamma \boldsymbol{r}_{ij} = \tilde{B}_j\boldsymbol{\lambda}_j,$$

and the Schwarz iteration (4.1) can be rewritten as

$$(\tilde{A}_\Gamma - \tilde{B}_i)\boldsymbol{\lambda}_i^{(n)} = M_\Gamma \boldsymbol{r}_{ij}^{(n)} + \boldsymbol{g}_\Gamma,$$
$$M_\Gamma \boldsymbol{r}_{ij}^{(n)} = \tilde{B}_j \boldsymbol{\lambda}_j^{(n-1)}.$$

We modify the second equation as suggested in [10] and [20] to the form

$$M_\Gamma \boldsymbol{r}_{ij}^{(n)} - \hat{p}\,\tilde{B}_i \boldsymbol{\lambda}_i^{(n)} = \tilde{B}_j \boldsymbol{\lambda}_j^{(n-1)} - \hat{p}\, M_\Gamma \boldsymbol{r}_{ji}^{(n-1)}$$

for $i, j \in \{1, 2\}$ and $i \neq j$. Here $0 \leq \hat{p} < 1$ is a parameter which we use for optimization. At convergence one recovers the original equations and therefore the fixed point of the iteration is the same as for the original method.

Remark 4.3. The above modification is shown in [20] to be equivalent (at the continuous level) to imposing

$$\left(\frac{1-\hat{p}}{1+\hat{p}}\,\mu + \frac{\partial}{\partial \boldsymbol{n}_i}\right) u_i^{(n)} = \left(\frac{1-\hat{p}}{1+\hat{p}}\,\mu + \frac{\partial}{\partial \boldsymbol{n}_i}\right) u_j^{(n-1)} \tag{4.7}$$

for $i, j \in \{1, 2\}$ and $i \neq j$. Note that if $\hat{p} = \frac{1-\sqrt{h}}{1+\sqrt{h}}$, then $\frac{1-\hat{p}}{1+\hat{p}}\mu \propto \frac{1}{\sqrt{h}}$ which is the right choice of parameter according to optimized Schwarz theory. We will see that this is exactly the right choice for $\hat{p}$ at the discrete level.

The analysis of this algorithm is possible using the framework established for the original method. We can eliminate the $\{r_{ij}\}$ as follows:

$$\begin{aligned}(\tilde{A}_\Gamma - \tilde{B}_i)\boldsymbol{\lambda}_i^{(n)} &= \hat{p}\tilde{B}_i\boldsymbol{\lambda}_i^{(n)} + \tilde{B}_j\boldsymbol{\lambda}_j^{(n-1)} - \hat{p}M_\Gamma \boldsymbol{r}_{ji}^{(n-1)} + \boldsymbol{g}_\Gamma \\ &= \hat{p}\tilde{B}_i\boldsymbol{\lambda}_i^{(n)} + \tilde{B}_j\boldsymbol{\lambda}_j^{(n-1)} - \hat{p}(\tilde{A}_\Gamma - \tilde{B}_j)\boldsymbol{\lambda}_j^{(n-1)} + (1+\hat{p})\boldsymbol{g}_\Gamma,\end{aligned}$$

which simplifies to

$$(\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_i)\boldsymbol{\lambda}_i^{(n)} = -(\hat{p}\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_j)\boldsymbol{\lambda}_j^{(n-1)} + (1+\hat{p})\boldsymbol{g}_\Gamma.$$

ALGORITHM 2. *Let* $\lambda_{h,1}^{(0)}, \lambda_{h,2}^{(0)} \in \Lambda_h$ *be two random initial guesses. Then for* $n = 1, 2, \ldots$ *find* $\{\lambda_{h,i}^{(n)}\}$ *such that*

$$\begin{aligned}(\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_1)\boldsymbol{\lambda}_1^{(n)} &= -(\hat{p}\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_2)\boldsymbol{\lambda}_2^{(n-1)} + (1+\hat{p})\boldsymbol{g}_\Gamma, \\ (\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_2)\boldsymbol{\lambda}_2^{(n)} &= -(\hat{p}\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_1)\boldsymbol{\lambda}_1^{(n-1)} + (1+\hat{p})\boldsymbol{g}_\Gamma.\end{aligned} \tag{4.8}$$

Since $\hat{p} < 1$, we can use Lemma 3.7 and conclude that the left-hand side is positive definite and therefore invertible. At convergence we have $(1-\hat{p})\tilde{A}_\Gamma(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2) = 0$, which implies $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \tilde{S}_\Gamma^{-1}\boldsymbol{g}_\Gamma$ if $\hat{p} \neq 1$.

Comparing to the original Schwarz method, Algorithm 1, we weakened the positive-definiteness of the left-hand side. This plays a key role in faster convergence. The optimized algorithm can be viewed as a different splitting of the Schur complement. More precisely we multiplied it by $(1+\hat{p})$ and this time a fraction of $\tilde{A}_\Gamma$, namely, $\hat{p}\tilde{A}_\Gamma$, has been put to the right-hand side.

We consider the error equation and we can proceed as before to obtain an iteration for $\boldsymbol{e}_i$ only,

$$\begin{aligned}(\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_i)\boldsymbol{e}_i^{(n)} = {}& (\hat{p}\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_j) \\ & \cdot(\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_j)^{-1} \cdot (\hat{p}\tilde{A}_\Gamma - (1+\hat{p})\tilde{B}_i)\boldsymbol{e}_i^{(n-2)}.\end{aligned}$$

With $\boldsymbol{\varphi}_i = (I-(1+\hat{p})C_i)\tilde{A}_\Gamma^{1/2}\boldsymbol{e}_i$, we have

$$\begin{aligned}\boldsymbol{\varphi}_i^{(n)} &= (\hat{p}I-(1+\hat{p})C_j)\cdot(I-(1+\hat{p})C_j)^{-1}\\ &\quad\cdot(\hat{p}I-(1+\hat{p})C_i)\cdot(I-(1+\hat{p})C_i)^{-1}\boldsymbol{\varphi}_i^{(n-2)}.\end{aligned}$$

Denoting by $\hat{D}_i := (\hat{p}I-(1+\hat{p})C_i)\cdot(I-(1+\hat{p})C_i)^{-1}$ and simplifying, we get

$$\hat{D}_i = I-(1-\hat{p})\left(I-(1+\hat{p})C_i\right)^{-1}, \tag{4.9}$$

which shows that $\hat{D}_i$ is symmetric. Therefore we have

$$\|\boldsymbol{\varphi}_i^{(n)}\|_2 \le \rho(\hat{D}_j)\,\rho(\hat{D}_i)\,\|\boldsymbol{\varphi}_i^{(n-2)}\|_2.$$

The estimate for the eigenvalues of $\hat{D}_i$ can be obtained as before. More precisely we have

$$\sigma(\hat{D}_i) = 1-\frac{1-\hat{p}}{1-(1+\hat{p})\,\sigma(C_i)}.$$

Recall that $\sigma(C_i)\in\left[\frac{1}{2}-c,\frac{1}{2}-C\frac{h}{\alpha H}\right]$ for $0<c<\frac{1}{2}$ and $C>0$ and independent of h, H. We can use $\hat{p}$ to optimize $\rho(\hat{D}_i)$. Following Remark 4.3, let us make the ansatz

$$\hat{p} = \frac{1-(\frac{h}{\alpha})^\gamma}{1+(\frac{h}{\alpha})^\gamma} < 1, \quad \gamma\in\mathbb{R}^+.$$

This implies that

$$1-\frac{1}{\frac{1}{2}+\frac{C}{H}(\frac{h}{\alpha})^{1-\gamma}} \le \sigma(\hat{D}_i) \le 1-\frac{1}{\frac{1}{2}+c(\frac{h}{\alpha})^{-\gamma}}. \tag{4.10}$$

Best performance is achieved if $\gamma := \frac{1}{2}$ which as $h\to 0$ leads to

$$-1+c_1\sqrt{\frac{h}{\alpha}} \le \sigma(\hat{D}_i) \le 1-c_2\sqrt{\frac{h}{\alpha}}. \tag{4.11}$$

Note that the iteration matrix, $\hat{D}_i$, is not positive definite anymore but it has a converging spectrum and the contraction factor is much better than the one in Algorithm 1. We summarize our results in the next theorem.

THEOREM 4.4. *There exists an $\alpha>0$ independent of H and h such that Algorithm 2 converges and the contraction factor is bounded by*

$$\rho \le 1-O(\sqrt{h}). \tag{4.12}$$

5. A multi-subdomain algorithm. We have introduced and analyzed a two-subdomain OSM so far. In this section we introduce a multi-subdomain algorithm for the IPH discretization. This algorithm is a natural generalization of the two-subdomain method. Often special care has to be taken in OSMs for classical FEM discretizations at cross-points, that is, a node which is shared by more than two subdomains; see [16, 17, 18]. This is not the case when we work with a DG discretization because subdomains communicate with each other only if they have an intersection of nonzero measure. Therefore the problem with cross-points does not arise, since a cross-point is of measure-zero, as at the continuous level.

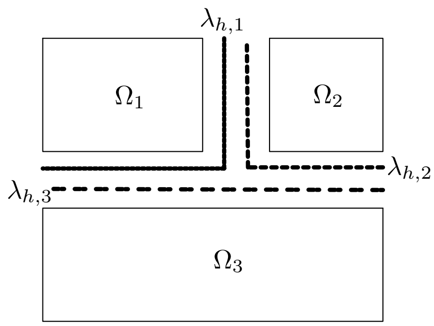

FIG. 3. *A multi-subdomain configuration with an interface variable, $\{\lambda_{h,i}\}$, assigned to each subdomain, Ω_i.*

Let us start defining the multi-subdomain geometry. We first partition the monodomain Ω into N_s subdomains such that the interface, Γ, between them is a subset of internal edges, $\mathcal{E}^0$. More precisely, we denote the subdomains by $\{\Omega_i\}_{i=1}^{N_s}$ and the interface between two subdomain by

$$\Gamma_{ij} := \partial\Omega_i \cap \partial\Omega_j \quad (i \neq j)$$

and the global interface by

$$\Gamma := \bigcup_{i \neq j} \Gamma_{ij} \subset \mathcal{E}^0.$$

Now the hybridizable formulation of IPH can be written as follows: find $(u_h, \lambda_h) \in V_h \times \Lambda_h$ such that

$$\tilde{a}((u_h, \lambda_h), (v, \varphi)) = (f, v)_{\mathcal{T}_h} \quad \forall (v, \varphi) \in V_h \times \Lambda_h, \tag{5.1}$$

where the bilinear form is defined as

$$\tilde{a}((u, \lambda), (v, \varphi)) := \tilde{a}_\Gamma(\lambda, \varphi) + \sum_{i=1}^{N_s} \left(\tilde{a}_i(u_i, v_i) + \tilde{a}_{i\Gamma}(u_i, \varphi) + \tilde{a}_{i\Gamma}(v_i, \lambda)\right). \tag{5.2}$$

The only modified bilinear form is $\tilde{a}_{i\Gamma}(\cdot,\cdot)$ since it acts now on $\partial\Omega_i \setminus \partial\Omega$, that is,

$$\tilde{a}_{i\Gamma}(u_i, \varphi) := \left\langle \frac{\partial u_i}{\partial \boldsymbol{n}_i} - \mu u_i, \varphi \right\rangle_{\partial\Omega_i \setminus \partial\Omega}. \tag{5.3}$$

Let us focus on two subdomains which share an interface, Γ_{ij}. We observe that there are two subproblems which are communicating through λ_h on Γ_{ij}. That is,

$$\begin{aligned} \tilde{a}_i(u_i, v_i) + \tilde{a}_{i\Gamma}(v_i, \lambda_h) &= (f, v_i)_{\mathcal{T}_i} \quad \forall v_i \in V_{h,i}, \\ \tilde{a}_j(u_j, v_j) + \tilde{a}_{j\Gamma}(v_j, \lambda_h) &= (f, v_j)_{\mathcal{T}_j} \quad \forall v_j \in V_{h,j}, \end{aligned}$$

and the continuity is imposed using

$$\lambda_h = \frac{1}{2\mu}\left(\mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}\right) + \frac{1}{2\mu}\left(\mu u_j - \frac{\partial u_j}{\partial \boldsymbol{n}_j}\right) \quad \text{on } \Gamma_{ij}. \tag{5.4}$$

Now we relax the constraint that λ_h is single-valued on Γ and allocate $\lambda_{h,i}$ to each subdomain Ω_i. Each $\lambda_{h,i}$ is defined on $\partial\Omega_i \setminus \partial\Omega$; for an example see Figure 3. We have therefore twice DOFs along Γ_{ij}. Therefore we should split the continuity equation (5.4) to provide two conditions, one for each $\lambda_{h,i}$. We use the same idea as in Algorithm 2 and relax the continuity equation in the same fashion:

$$\frac{1}{1+\hat{p}}\lambda_{h,i} + \frac{\hat{p}}{1+\hat{p}}\lambda_{h,j} = \frac{1}{2\mu}\left(\mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}\right) + \frac{1}{2\mu}\left(\mu u_j - \frac{\partial u_j}{\partial \boldsymbol{n}_j}\right), \quad (i \neq j).$$

Here $\hat{p}$ is a parameter which is used for optimization purposes. This suggests the following iterative method to find the pairs $\left\{(u_i, \lambda_{h,i})\right\}_{i=1}^{N_s}$ in parallel.

ALGORITHM 3. *Let* $\left\{(u_i^{(0)}, \lambda_{h,i}^{(0)})\right\}_{i=1}^{N_s}$ *be a set of initial guesses for all subdomains. Then for* $n = 1, 2, \ldots$ *find* $\left\{(u_i^{(n)}, \lambda_{h,i}^{(n)})\right\}_{i=1}^{N_s}$ *such that*

$$\tilde{a}_i(u_i^{(n)}, v_i) + \tilde{a}_{i\Gamma}(v_i, \lambda_{h,i}^{(n)}) \;=\; (f, v_i)_{\mathcal{T}_i} \quad \forall v_i \in V_{h,i}, \tag{5.5}$$

and the continuity condition on Γ_{ij} *reads*

$$\frac{1}{1+\hat{p}}\lambda_{h,i}^{(n)} - \frac{1}{2\mu}\left(\mu u_i - \frac{\partial u_i}{\partial \boldsymbol{n}_i}\right)^{(n)} = -\frac{\hat{p}}{1+\hat{p}}\lambda_{h,j}^{(n-1)} + \frac{1}{2\mu}\left(\mu u_j - \frac{\partial u_j}{\partial \boldsymbol{n}_j}\right)^{(n-1)}. \tag{5.6}$$

At convergence we obtain $(1-\hat{p})(\lambda_{h,i} - \lambda_{h,j}) = 0$. Therefore if $\hat{p} \neq 1$, we recover that λ_h is single valued.

Remark 5.1. We can make an ansatz for the optimal choice of $\hat{p}$ similar to the two-subdomain case. The transmission condition (5.4) can be viewed as a Robin transmission condition at the continuous level. The Robin parameter is $\mu^\star := \frac{1-\hat{p}}{1+\hat{p}}\mu$. In order to converge fast we should set $\mu^\star = O(h^{-1/2})$. This corresponds to the choice $\hat{p} := \frac{1-\sqrt{h}}{1+\sqrt{h}} < 1$.

5.1. OSM as a preconditioner. We show now how one can use OSM as a preconditioner for a Krylov subspace method. We start by writing Algorithm 3 at the algebraic level. We first partition the DOFs associated with $u_h \in V_h$ into

$$\boldsymbol{u} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{N_s})^\top.$$

Then we form DOFs associated to the interface unknowns $\{\lambda_{h,i}\}_{i=1}^{N_s}$ by

$$\boldsymbol{\ell} := (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_{N_s})^\top$$

and define the augmented DOFs by $\boldsymbol{w} := (\boldsymbol{u}, \boldsymbol{\ell})^\top$.

Algorithm 3 can be written at the algebraic level as

$$\underbrace{\begin{bmatrix} K_{uu} & K_{u\ell} \\ K_{\ell u} & K_{\ell\ell} \end{bmatrix}}_{K} \boldsymbol{w}^{(n)} = \underbrace{\begin{bmatrix} 0 & 0 \\ L_{\ell u} & L_{\ell\ell} \end{bmatrix}}_{L} \boldsymbol{w}^{(n-1)} + \underbrace{\begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix}}_{\boldsymbol{g}}. \tag{5.7}$$

Note that the left-hand-side matrix K consists of block matrices communicating only with each pair $(u_{h,i}, \lambda_{h,i})$. Therefore we can "invert" subdomain blocks independently and in parallel. This gives a parallel preconditioner for a Krylov subspace method applied to the system $(K - L)\boldsymbol{w} = \boldsymbol{g}$.

TABLE 1
Minimum and maximum eigenvalues of C_i.

$\sqrt{N}$	6	13	26	55	112	225
$\sigma_{\min}$	0.295	0.288	0.286	0.286	0.286	0.286
$\sigma_{\max}$	0.335	0.415	0.457	0.478	0.489	0.494

Since the stationary iterates (5.7) converge with the contraction factor $\rho \leq 1 - O(\sqrt{h})$, we expect that a preconditioned Krylov subspace method achieves another square root in the contraction factor, that is, $\rho \leq 1 - O(h^{1/4})$. This is observed in the numerical experiments. Therefore this is a more attractive method compared to the CG method with an additive Schwarz preconditioner which has the contraction factor $\rho \leq 1 - O(\sqrt{h})$.

6. Numerical experiments. We perform numerical experiments on the model problem

$$(6.1)\qquad \begin{aligned} (\eta - \Delta)u &= f \qquad \text{in } \Omega,\\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned}$$

where $\eta = 1$ and Ω is either a unit square, i.e., $(0,1)^2$, or an L-shaped (nonconvex) domain. The interface is such that it does not cut through any element; therefore $\Gamma \subset \mathcal{E}$. We use $\mathbb{P}^1$ elements and $\alpha = c(k+1)(k+2)$, where $c > 0$ is a constant independent of h and $k = 1$ (polynomial degree). The algorithms are implemented using a FORTRAN90 library for DG methods called `GDG90`. The codes are accessible at http://unige.ch/~hajian/gdg90.

6.1. Minimum and maximum eigenvalues of B_i. Before performing convergence experiments on Algorithms 1 and 2, let us validate numerically the asymptotic behavior of the minimum and maximum eigenvalues of the operator B_i, i.e., inequality (3.1). To do so, we should measure the minimum and maximum eigenvalues of $C_i := \tilde{A}_\Gamma^{-1/2} B_i \tilde{A}_\Gamma^{-1/2}$. We generate a sequence of quasi-uniform triangulations and construct the operators B_i and $\tilde{A}_\Gamma$ for each triangulation. We denote the size of each operator by N, i.e., $B_i \in \mathbb{R}^{N\times N}$. We have $1/h \propto \sqrt{N}$ as h goes to zero.

According to (4.4), the minimum eigenvalue of C_i is bounded from below independently of the mesh size. This can be seen from Table 1. For the maximum eigenvalues of C_i, observe that $\sigma_{\max}$ is less than $\frac{1}{2}$ and is increasing. In order to see the growth rate we plot $\frac{1}{2} - \sigma_{\max}$ in Figure 4 which decreases like $1/\sqrt{N} = O(h)$ as N goes to infinity. This is in agreement with (4.5).

6.2. Two-subdomain case. In this section we compare the contraction factor of the two Schwarz algorithms with respect to h-dependency. We perform both algorithms on a sequence of unstructured meshes. We measure the number of iterations required to reduce the relative error to $tol :=$ 1E-10 while refining the mesh, that is,

$$\|u_h^{(n)} - u_h\|_0 \leq tol\,\|f\|_0.$$

This level of accuracy is not necessary in practice since the error between the exact and approximate solutions, $\|u - u_h\|_0$, is much bigger and one usually can terminate the iteration after reaching the accuracy level of the method. The domain is partitioned into two by a nonstraight interface; see Figure 1 (left).

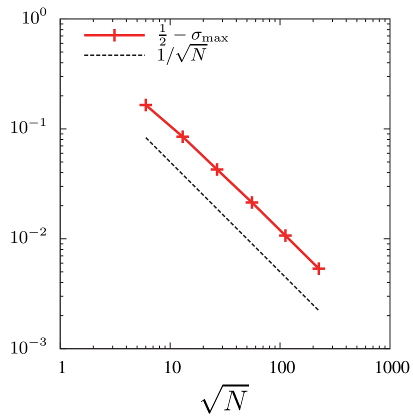

FIG. 4. *Behavior of* $\left(\frac{1}{2} - \sigma_{\max}\right)$ *versus total number of unknowns,* N*.*

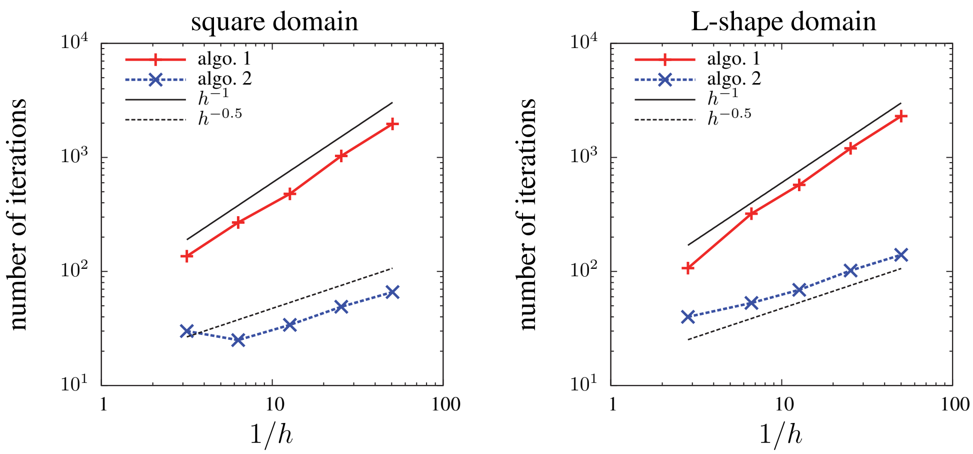

FIG. 5. *Convergence of Schwarz methods on a square domain (left) and L-shaped domain (right).*

As we see in the Figure 5 (left), on a square domain the number of iterations for Algorithm 1 grows like $1/h$, which is equivalent to $\rho \leq 1 - O(h)$, while for Algorithm 2 it behaves like $1/\sqrt{h}$, or in other words we have $\rho \leq 1 - O(\sqrt{h})$, which illustrates well our analysis. This is the case for the L-shaped domain too; see Figure 5 (right).

6.3. Multi-subdomains case. We now show some numerical results on the multi-subdomain algorithm. The subdomains are formed by a coarse triangulation of the domain which we call $\mathcal{T}_H$. We consider a nested fine mesh and therefore $\mathcal{T}_H \subset \mathcal{T}_h$. An example is given in Figure 1 (right). We consider here *four* subdomains which share a cross-point, and similarly to the two subdomain case we measure the number of iterations necessary to reach the desired tolerance. We observe in Table 2 that the contraction factor asymptotically is $\rho = 1 - O(\sqrt{h})$, i.e., $82/57 \approx 1.43$ or $117/82 \approx 1.42$, which are close to $\sqrt{2}$.

6.4. OSM as a preconditioner. We use now the OSM as a preconditioner for GMRES with the tolerance *tol* := 1E-6. In order to provide a qualitative comparison we also consider the widely used CG method with a one-level additive Schwarz preconditioner applied to the original system (1.2). We consider 16 subdomains

TABLE 2
Convergence of OSM for four subdomains.

Mesh size	h_0	$h_0/2$	$h_0/4$	$h_0/8$	$h_0/16$
# iterations	25	35	57	82	117

TABLE 3
Number of iterations required by OSM-GMRES and PCG to reach the desired tolerance.

Mesh size	h_0	$h_0/2$	$h_0/4$	$h_0/8$	$h_0/16$
OSM-GMRES	20	52	60	72	87
PCG	14	38	55	104	154

illustrated in Figure 1 (right). We observe in Table 3 that the number of iterations for OSM-GMRES grows like $O(h^{-1/4})$. This is because Krylov methods benefit often from another square root in their contraction factor compared to the stationary iteration method. Therefore the contraction factor of OSM-GMRES is $\rho = 1 - O(h^{1/4})$, i.e., $72/60 \approx 1.2, 87/72 \approx 1.2$, which are close to $2^{1/4}$. For the preconditioned (additive Schwarz) CG method, we have $\rho = 1 - O(\sqrt{h})$.

We would like to comment on the size of the augmented system. In case of mesh size $h_0/16$ we have 19,032 DOFs for the primal variable u_h and 1,296 DOFs for the interface unknowns. Therefore the augmented system is very little changed in size compared to the original system.

7. Conclusion. We have presented and analyzed classical and optimized Schwarz methods for IPH discretizations. The interesting fact is that both use Robin transmission conditions, but we proved that for an arbitrary two-subdomain decomposition the classical Schwarz algorithm has a convergence factor $1-O(h)$, while the optimized one has a contraction factor $1 - O(\sqrt{h})$. This is because the IPH discretization imposes a bad choice of the Robin parameter on the method. We then generalized the definition of the algorithms to the multi-subdomain case and showed by numerical experiments that our theoretical results still hold. We finally illustrated the potential benefit that one obtains using OSM as a preconditioner compared to PCG.

Appendix. In this part we provide some proofs regarding the extension by zero operator, $\boldsymbol{\theta}_i(\cdot)$. First we recall inverse and mass matrix inequalities; see [25, Appendix B] and references therein. All constants are independent of h. Let $w \in \mathbb{P}^1(K)$, where K is a simplex in $\mathbb{R}^d$. Then the inverse inequality

$$\|\nabla w\|_K \leq \frac{c}{h}\|w\|_K \tag{A.1}$$

holds. Let $\boldsymbol{w}$ be the DOFs of w and M_d be the corresponding mass matrix. Then we have

$$c_1\, h^d\, \boldsymbol{w}^\top \boldsymbol{w} \leq \boldsymbol{w}^\top M_d \boldsymbol{w} \leq c_2\, h^d\, \boldsymbol{w}^\top \boldsymbol{w}.$$

LEMMA A.1. *Let $\varphi \in \Lambda_h$ and $\boldsymbol{\theta}_i(\varphi)$ be its extension by zero operator into Ω_i. For an element K which shares an edge with the interface, we have*

$$\|\nabla \boldsymbol{\theta}_i(\varphi)\|_K^2 \leq C_1 h^{-1} \|\varphi\|_e^2, \tag{A.2}$$

$$\|\boldsymbol{\theta}_i(\varphi)\|_K^2 \leq C_2\, h \|\varphi\|_e^2. \tag{A.3}$$

Proof. Let $\boldsymbol{\varphi}_e := (\varphi_1, \varphi_2)$ be the DOFs of φ on the edge shared with the interface. Moreover let $w = \boldsymbol{\theta}_i(\varphi)|_K$. Then we have $\boldsymbol{w} = (\varphi_1, \varphi_2, 0)$. For the first inequality we invoke the inverse inequality. Assuming the mesh is quasi-uniform, i.e., $h_e \approx h_K \approx h$, we get

$$\begin{aligned}\|\nabla w\|_K^2 \le \frac{c^2}{h^2}\|w\|_K^2 &\le c_1 h^{d-2}(\varphi_1^2 + \varphi_2^2 + 0)\\ &\le c_2 h^{d-2} h_e^{-(d-1)} \boldsymbol{\varphi}_e^\top M_{d-1} \boldsymbol{\varphi}_e\\ &\le c_3 h^{-1} \boldsymbol{\varphi}_e^\top M_{d-1} \boldsymbol{\varphi}_e\\ &= c_3 h^{-1} \|\varphi\|_e^2.\end{aligned}$$

The proof for the second inequality follows the same steps. ∎

LEMMA A.2. *Let $\varphi \in \Lambda_h$ and $\boldsymbol{\theta}_i(\varphi)$ be its extension by zero operator into Ω_i. Then*

$$\|[\![\boldsymbol{\theta}_i(\varphi)]\!]\|_{\mathcal{E}_i}^2 \le C\|\varphi\|_\Gamma^2,$$

where $C \ge 1$.

Proof. We start by those edges which are part of the interface (see Figure 2), e.g., e_1 and e_3. We have

$$\sum_{e\in\Gamma} \|[\![\boldsymbol{\theta}_i(\varphi)]\!]\|_e^2 = \sum_{e\in\Gamma} \|\boldsymbol{\theta}_i(\varphi)\|_e^2 = \sum_{e\in\Gamma} \|\varphi\|_e^2 = \|\varphi\|_\Gamma^2,$$

which shows already that $C \ge 1$. Consider those edges $e \in \mathcal{E}_i$ that are not on the interface but belong to an element which shares an edge with the interface, e.g., $e^* := \partial K_1 \cap \partial K_2$ in Figure 2. Let $\boldsymbol{\varphi}_e := (\varphi_1, \varphi_2)$ be the DOFs of φ on e_2 and assume φ_2 is the DOF which is also located on e^*. Then we have

$$\|[\![\boldsymbol{\theta}_i(\varphi)]\!]\|_{e^*}^2 = (\varphi_2, 0) M_{d-1} (\varphi_2, 0)^\top \le c h_{e^*}^{d-1} \varphi_2^2 \le c h_{e^*}^{d-1}(\varphi_2^2 + \varphi_1^2) \le c_1 \|\varphi\|_e^2,$$

where we again used the quasi-uniformity of the mesh ($h_e \approx h \approx h_{e^*}$). The other case would be that K_1 and K_3 share an edge, for which we can use the same argument. For other edges $[\![\boldsymbol{\theta}_i(\varphi)]\!]$ is simply zero. ∎

REFERENCES

[1] P. F. ANTONIETTI AND B. AYUSO, *Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems*, ESAIM Math. Model. Numer. Anal., 42 (2008), pp. 443–469.

[2] P. F. ANTONIETTI AND P. HOUSTON, *A class of domain decomposition preconditioners for hp-discontinuous Galerkin finite element methods*, J. Sci. Comput., 46 (2011), pp. 124–149.

[3] 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 (2002), pp. 1749–1779.

[4] S. C. BRENNER, *The condition number of the Schur complement in domain decomposition*, Numer. Math., 83 (1999), pp. 187–203.

[5] S. C. BRENNER, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal., 41 (2003), pp. 306–324.

[6] P. CASTILLO, *Performance of discontinuous Galerkin methods for elliptic PDEs*, SIAM J. Sci. Comput., 24 (2002), pp. 524–547.

[7] B. COCKBURN, J. GOPALAKRISHNAN, AND R. LAZAROV, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.

[8] M. DISCACCIATI, *An Operator-Splitting Approach to Nonoverlapping Domain Decomposition Methods*, Technical report, EPFL, Lausanne, Switzerland, 2004.

[9] M. DRYJA, J. GALVIS, AND M. SARKIS, *BDDC methods for discontinuous galerkin discretization of elliptic problems*, J. Complexity, 23 (2007), pp. 715–739.

[10] M. El Bouajaji, V. Dolean, M. J. Gander, S. Lanteri, and R. Perrussel, *DG discretization of optimized Schwarz methods for Maxwell's equations*, in Domain Decomposition Methods in Science and Engineering XXI, Lect. Notes Comput. Sci. Eng. 98, Springer, New York, 2014, pp. 217–225.

[11] R. E. Ewing, J. Wang, and Y. Yang, *A stabilized discontinuous finite element method for elliptic problems*, Numer. Linear Algebra Appl., 10 (2003), pp. 83–104.

[12] 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), pp. 1343–1365.

[13] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.

[14] M. J. Gander, *Schwarz methods over the course of time*, Electron. Trans. Numer. Anal., 31 (2008), pp. 228–255.

[15] M. J. Gander and S. Hajian, *Block Jacobi for discontinuous Galerkin discretizations: No ordinary Schwarz methods*, in Domain Decomposition Methods in Science and Engineering XXI, Lect. Notes Comput. Sci. Eng., Springer, New York, 2013, pp. 305–313.

[16] M. J. Gander and F. Kwok, *Best Robin parameters for optimized Schwarz methods at cross points*, SIAM J. Sci. Comput., 34 (2012), pp. A1849–A1879.

[17] M. J. Gander and F. Kwok, *On the applicability of Lions energy estimates in the analysis of discrete optimized Schwarz methods with cross points*, in Domain Decomposition Methods in Science and Engineering XX, Lect. notes Comput. Sci. Eng. 91, Springer, New York, 2013, pp. 475–483.

[18] M. J. Gander and K. Santugini, *Cross-points in domain decomposition methods with a finite element discretization*, arXiv:1404.4698 [math.NA], 2014.

[19] C. J. Gittelson, R. Hiptmair, and I. Perugia, *Plane wave discontinuous Galerkin methods: Analysis of the h-version*, ESAIM Math. Model. Numer. Anal., 43 (2009), pp. 297–331.

[20] S. Hajian, *An optimized Schwarz algorithm for discontinuous Galerkin methods*, in Domain Decomposition Methods in Science and Engineering XXII, Lect. Notes Comput. Sci. Eng. 96, Springer, New York, 2014.

[21] C. Lehrenfeld, *Hybrid Discontinuous Galerkin Methods for Incompressible Flow Problems*, Master's thesis, RWTH Aachen, 2010.

[22] S. H. Lui, *A Lions non-overlapping domain decomposition method for domains with an arbitrary interface*, IMA J. Numer. Anal., 29 (2009), pp. 332–349.

[23] L. Qin and X. Xu, *On a parallel Robin-type nonoverlapping domain decomposition method*, SIAM J. Numer. Anal., 44 (2006), pp. 2539–2558.

[24] J. Schberl and C. Lehrenfeld, *Domain Decomposition preconditioning for high order hybrid discontinuous Galerkin methods on tetrahedral meshes*, in Advanced Finite Element Methods and Applications, T. Apel and O. Steinbach, eds., Lect. Notes Appl. Comput. Mech. 66, Springer, Berlin, 2013, pp. 27–56.

[25] A. Toselli and O. Widlund, *Domain decomposition Methods—Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005.

[26] T. Warburton and J. S. Hesthaven, *On the constants in hp-finite element trace inverse inequalities*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 2765–2773.