SIAM J. NUMER. ANAL.
Vol. 44, No. 4, pp. 1612–1638

A DOMAIN DECOMPOSITION METHOD BASED ON WEIGHTED INTERIOR PENALTIES FOR ADVECTION-DIFFUSION-REACTION PROBLEMS*

ERIK BURMAN† AND PAOLO ZUNINO‡

Abstract. We propose a domain decomposition method for advection-diffusion-reaction equations based on Nitsche's transmission conditions. The advection-dominated case is stabilized using a continuous interior penalty approach based on the jumps in the gradient over element boundaries. We prove the convergence of the finite element solutions of the discrete problem to the exact solution and propose a parallelizable iterative method. The convergence of the resulting domain decomposition method is proved, and this result holds true uniformly with respect to the diffusion parameter. The numerical scheme that we propose here can thus be applied straightforwardly to diffusion-dominated, advection-dominated, and hyperbolic problems. Some numerical examples are presented in different flow regimes showing the influence of the stabilization parameter on the performance of the iterative method, and we compare our method with some other domain decomposition techniques for advection-diffusion equations.

Key words. advection-diffusion problem, interior penalty, finite element approximation, domain decomposition, iterative methods, discontinuous coefficients

AMS subject classifications. 65N30, 65N12, 35L50, 65N55

DOI. 10.1137/050634736

1. Introduction. The solution of large computational problems calls for efficient linear solvers. Domain decomposition has proved to be an attractive way to allow for parallel solving of large problems. A formulation for domain decomposition using a generalization of Nitsche's method for weak boundary conditions has been considered, for instance, by Becker, Hansbo, and Stenberg [2, 24] and by Heinrich and Pietsch [16] for the Poisson problem. This formulation was then generalized to the case of advection-diffusion problems by Toselli [26] using SUPG-type stabilization and more recently by Burman [5]. In this last case, continuous interior penalty stabilization was used to make the method stable in all flow regimes. The interior penalty finite element method for continuous approximation spaces was introduced by Douglas and Dupont [12] and analyzed by Burman and Hansbo in [7] and by Burman in [5].

In this paper we will give a detailed analysis of the domain decomposition method using Nitsche's method. In particular we consider a fully parallel iterative splitting method for advection-diffusion-reaction problems, and we prove its convergence. The present result also automatically carries over to discontinuous Galerkin interior penalty formulations of advection-diffusion problems. Overlapping domain decomposition methods for discontinuous Galerkin methods was considered by Lasser and Toselli [19] and substructuring iterative methods for domain decomposition using SUPG-type stabilized continuous approximation was considered by Rapin and Lube [23]. For an overview of results on domain decomposition for nonsymmetric problems, see Quarteroni and Valli [22] or Toselli and Widlund [27] and the references therein.

*Received by the editors June 29, 2005; accepted for publication (in revised form) March 23, 2006; published electronically August 7, 2006.
http://www.siam.org/journals/sinum/44-4/63473.html

†Institut d'analyse et calcul scientifique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland (Erik.Burman@epfl.ch).
‡MOX, Dipartimento di Matematica, Politecnico di Milano, P.zza L. Da Vinci 32, Milano, Italy (paolo.zunino@polimi.it).

The advantages of the method proposed in this paper are to allow for continuous and discontinuous approximation with uniform stability properties with respect to the Péclet number. The discontinuous formulation naturally leads to an iterative method and allows for conservation locally in each subdomain. The continuous approximation, on the other hand, is better suited to handle different diffusive regimes since the interior penalty stabilization parameter is independent of the diffusion parameter. Numerical tests show that the proposed method is robust with respect to varying coefficients. As a model problem we propose the advection-diffusion-reaction equation

$$\left\{ \begin{aligned} \beta\cdot\nabla u + \sigma u - \nabla\cdot\varepsilon\nabla u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \right. \tag{1.1}$$

where Ω is a bounded open connected subset of $\mathbb{R}^d$ with a Lipschitz boundary $\partial\Omega$, $d = 2$ or 3 is the space dimension, $\beta \in [W^{1,\infty}(\Omega)]^d$ is a velocity field, $\varepsilon \in L^\infty(\Omega)$, $\varepsilon > 0$, is a diffusion coefficient, and $\sigma > 0$ is the reaction coefficient, $f \in L^2(\Omega)$. The analysis extends to the case $\varepsilon = 0$ in the obvious way if the boundary conditions of the continuous problem are modified and β is such that the problem remains well-posed. We assume that the following coercivity condition holds:

$$\sigma - \frac{1}{2}\nabla\cdot\beta \geq \sigma_0 > 0. \tag{1.2}$$

We define the associated parameter σ_1 by

$$\sigma_1 := \text{ess sup}_{x\in\Omega}\frac{|\sigma - \nabla\cdot\beta|^2}{\sigma_0}.$$

Consider a decomposition of the domain Ω into the disjoint subdomains Ω_i, $i = 1,\dots,N$, with boundaries $\partial\Omega_i$ and with corresponding shape regular disjoint triangulations $\mathcal{T}_{h,i}$, such that $\mathcal{T}_h = \cup_{i=1}^N \mathcal{T}_{h,i} = \cup_{i=1}^N \bar{\Omega}_i = \bar{\Omega}$. Note that we do not suppose that neighboring meshes are conforming over the intersubdomain boundary. The set of interior faces of each triangulation $\mathcal{T}_{h,i}$ will be denoted by $\mathcal{F}_i$. On each triangulation we define a finite element space $V_{h,k,i}$ associated with the subdomain Ω_i,

$$V_{h,k,i} := \{v_h : v_h \in H^1(\Omega_i); v_h|_K \in P_k(K)\ \forall K \in \mathcal{T}_{h,i}\},$$

where $P_k(K)$ denotes the space of polynomials of degree $\leq k$ on K and we let $V_h = \sum_{i=1}^N V_{h,k,i}$. For every function $v_h \in V_h$ we introduce the restriction to subdomain Ω_i, $v_{h,i} = v_h|_{\Omega_i}$. To each subdomain boundary we associate the outward-oriented normal n_i. We will always assume that the solution is sufficiently smooth, i.e., $u \in H^1(\Omega) \cap (\cup_{i=1}^N H^2(\Omega_i))$, and we will assume (weak) continuity of fluxes between subdomains. Typically the diffusion parameter ε may be discontinuous over some subdomain interface, provided the interface is smooth. Let h_K denote the diameter of an element K, and ϱ_K the radius of the largest inscribed ball in K. We henceforth assume that for all meshes $\mathcal{T}_{h,i}$ there holds

$$c_{\mathcal{T}} \leq \max_{K\in\mathcal{T}_{h,i}} \frac{h_K}{\varrho_K} \tag{1.3}$$

with the same positive parameter $c_{\mathcal{T}}$. We introduce a mesh parameter function $\tilde{h}(x)|_K = h_K$ and let $h = \max_{K\in\mathcal{T}_{h,i}} h_K$. Moreover we shall assume that there exists a constant $\rho > 1$ such that for all elements K in $\mathcal{T}_{h,i}$, $i = 1, \ldots, N$, we have

$$\max_{K' \in \mathcal{N}(K)} h_{K'} \leq \rho \min_{K' \in \mathcal{N}(K)} h_{K'}, \tag{1.4}$$

where $\mathcal{N}(K)$ is the set of elements K' such that $\bar{K} \cap \bar{K}' \neq \emptyset$. Property (1.4) is a local quasi-uniformity property of the mesh. The jump $[x]|_E$ of a quantity x over a face E will be defined by $[x(\xi)]|_E = \lim_{\delta \to 0}(x(\xi - n_E\delta) - x(\xi + n_E\delta))$, where $\xi \in E$ and n_E denotes a normal vector to the face E for interior faces where the normal is fixed but arbitrary, while for faces on a subdomain boundary $E \in \partial\Omega_i$ the normal is outward oriented with respect to the subdomain Ω_i and denoted n_i. Subscripts will be omitted when there is no ambiguity. For faces such that $E \cap \partial\Omega \neq \emptyset$ we set $[x(\xi)]|_E \equiv \lim_{\delta \to 0} x(\xi - n_E\delta)$. By $\{x(\xi)\}|_E$ we denote the average value of x over face E, $\{x(\xi)\}|_E = \lim_{\delta \to 0} \frac{1}{2}(x(\xi - n_E\delta) + x(\xi + n_E\delta))$. We will also use the weighted average $\{x(\xi)\}_w|_E = \lim_{\delta \to 0}(w^- x(\xi - n_E\delta) + w^+ x(\xi + n_E\delta))$, where w^- and w^+ are two positive weights such that $w_- + w_+ = 1$, and for faces on the boundary $\partial\Omega$ we define $\{x(\xi)\}|_E = \{x(\xi)\}_w|_E = \lim_{\delta \to 0} 2x(\xi - n_E\delta)$. Furthermore we will use the notation $(x, y)_X = \int_X x \cdot y \, \mathrm{dx}$, $\langle x, y \rangle_{\partial X} = \int_{\partial X} x \cdot y \, \mathrm{ds}$ with the elementwise counterparts $(x, y)_{X,h} = \sum_{K \in X} \int_K x \cdot y \, \mathrm{dx}$ and $\langle x, y \rangle_{\partial X,h} = \sum_{E \in \partial X} \int_E x \cdot y \, \mathrm{ds}$. Let $\|x\|_X = (x, x)_X^{\frac{1}{2}}$ denote the L^2-norm over X with the elementwise counterpart $\|x\|_{X,h} = (x, x)_{X,h}^{\frac{1}{2}}$. The norm of the space $H^i(X)$ will be denoted $\|x\|_{i,X}$ with $i = 0, 1, 2, \ldots$. The notations $\|x\|_X$ and $\|x\|_{0,X}$ are equivalent. The latter will be used only where it is more appropriate. For other functional spaces the notation will be made completely explicit. We will use c and C to denote generic positive constants independent of h_K but not necessarily of the local mesh geometry.

2. A domain decomposition method based on interior penalties. In this section we will show how domain decomposition using Nitsche's method leads to a continuous/discontinuous Galerkin-type penalty method in a natural way. The approximation is chosen to be continuous on each subdomain. We consider problem (1.1) on Ω and by taking V_h as trial and test space we propose the finite element method: find $u_h \in V_h$ such that

$$A(u_h, v_h) + J(u_h, v_h) + B(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h, \tag{2.1}$$

where

$$A(u_h, v_h) := \sum_{i=1}^{N} \Big(((\sigma - \nabla \cdot \beta)u_h, v_h)_{\Omega_i} + (\varepsilon \nabla u_h, \nabla v_h)_{\Omega_i} - (u_h, \beta \cdot \nabla v_h)_{\Omega_i} \Big),$$

$$J(u_h, v_h) := \sum_{i=1}^{N} \sum_{E \in \mathcal{F}_i} \left\langle \tilde{\gamma}_{1,i}(h_E) \|\beta \cdot n\|_{L^\infty(E)} [\nabla u_h \cdot n], [\nabla v_h \cdot n] \ \right\rangle_E,$$

$$B(u_h, v_h) := \sum_{i=1}^{N} \Bigg(\left\langle \beta \cdot n_i^+ u_h, [v_h] \right\rangle_{\partial\Omega_i} - \frac{1}{2} \left\langle \{\varepsilon \nabla u_h \cdot n_i\}_w, [v_h] \right\rangle_{\partial\Omega_i}$$
$$- \frac{1}{2} \left\langle \{\varepsilon \nabla v_h \cdot n_i\}_w, [u_h] \right\rangle_{\partial\Omega_i} + \left\langle \frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} [u_h], [v_h] \right\rangle_{\partial\Omega_i} \Bigg),$$

and $\beta \cdot n_i^{\pm} := \frac{1}{2}(|\beta \cdot n_i| \pm \beta \cdot n_i)$. The discretization of the advection term corresponds to the standard upwind flux after integration by parts. Note that the bilinear form A corresponds to a standard Galerkin formulation in each subdomain, supplemented with boundary terms on the inner and outer boundaries that appear naturally in the formulation to assure coercivity or consistency. We observe that terms associated with nonhomogeneous boundary data do not appear since we consider $u = 0$ on $\partial\Omega$. The interior penalty stabilization term has been decomposed into one term controlling the jumps in the gradient over *interior* faces of each subdomain Ω_i, that is, $J(u_h, v_h)$, and the terms controlling the jump of the solution over interior boundaries of neighboring subdomains, the upwind flux term and the penalty term $\langle(\gamma_{bc}\{\varepsilon\}_w/\tilde{h})[u_h],[v_h]\rangle_{\partial\Omega_i}$. The stabilization parameter $\tilde{\gamma}_{1,i}(h_E) = \gamma_{ip,i}h_E^2$ depends only on the mesh geometry of the subdomain triangulation $\mathcal{T}_{h,i}$.

Remark 2.1. If the triangulation of each subdomain consists of a single triangle, then the formulation (2.1) is equivalent to a standard interior penalty discontinuous Galerkin method for (1.1). This follows immediately by noting that the interior penalty term on the gradient jumps vanishes since there are no interior faces in the subdomains.

Remark 2.2. Recalling the framework for discontinuous Galerkin methods based on interior penalties by Arnold et al. [1], we observe that the definition of the coupling term $B(u_h, v_h)$ can be made more general by introducing a parameter s that allows us to switch between a symmetric and a nonsymmetric version. Precisely, we consider

$$B(u_h,v_h) := \sum_{i=1}^{N}\Bigg(\big\langle \beta\cdot n_i^{+}u_h,[v_h]\big\rangle_{\partial\Omega_i} - \frac{1}{2}\left\langle\{\varepsilon\nabla u_h\cdot n_i\}_w,[v_h]\right\rangle_{\partial\Omega_i} - \frac{s}{2}\left\langle\{\varepsilon\nabla v_h\cdot n_i\}_w,[u_h]\right\rangle_{\partial\Omega_i} + \left\langle\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}}[u_h],[v_h]\right\rangle_{\partial\Omega_i}\Bigg),$$

where the symmetric and the nonsymmetric cases are obtained by $s = 1$ and $s = -1$, respectively. In this work we mainly consider $s = 1$, but for comparison the nonsymmetric case will be addressed in section 4.

2.1. A priori error estimate. In this section we will prove that the finite element solution obtained from formulation (2.1) converges to the exact solution of (1.1). The a priori error estimate is proved using the techniques from [2] for the Nitsche matching conditions combined with the technique of [5] for the interior penalty stabilization. The main idea behind the stabilization based on the jump in the gradient between adjacent elements is to introduce a least squares control over the part of the convective derivative that is not in the finite element space. A key result is the following lemma. For a proof of the underlying approximation result between discrete spaces we refer to [18], and for a proof in the context of interior penalty stabilization we refer to [6]. First we define the Oswald quasi-interpolant π_h^* (see [17]).

DEFINITION 2.3. *For each node x_i, let n_i be the number of elements containing x_i as a node. We define a quasi-interpolant π_h^* of degree k by*

$$\pi_h^* v(x_i) := \frac{1}{n_i}\sum_{\{K\,:\,x_i\in K\}} v_{|K}(x_i) \quad \forall v \in \{v : v|_K \in P_k(K)\}.$$

THEOREM 2.4 (stability). *Let $\beta_h \in [V_{h,1,i}]^d$ be the Lagrange interpolant of β and let $u_h \in V_{h,k,i}$. Then there exists a constant $\gamma_{ip,i} \geq c_0 > 0$, depending only on the* *local mesh geometry, such that*

$$\|\tilde{h}^{\frac{1}{2}}(\beta_h \cdot \nabla u_h - \pi_h^*(\beta_h \cdot \nabla u_h))\|_{\Omega_i}^2 \leq J_i(u_h, u_h)$$

with

$$J_i(u_h, u_h) = \sum_{E \in \mathcal{F}_i} \int_E \gamma_{ip,i} h_E^2 \|\beta_h \cdot n\|_{L^\infty(E)} [\nabla u_h]^2 \; ds. \tag{2.2}$$

Remark 2.5. Clearly then $\|\tilde{h}^{\frac{1}{2}}(\beta_h \cdot \nabla u_h - \pi_h^*(\beta_h \cdot \nabla u_h))\|_{\Omega_i}^2 \leq J(u_h, u_h)$ since $\|\beta_h \cdot n\|_{L^\infty(E)} \leq \|\beta \cdot n\|_{L^\infty(E)}$.

We define a triple norm on each subdomain as

$$|||w_h|||_i^2 = \|\sigma_0^{\frac{1}{2}} w_h\|_{\Omega_i}^2 + \|\varepsilon^{\frac{1}{2}} \nabla w_h\|_{\Omega_i}^2 + J_i(w_h, w_h) \tag{2.3}$$

and the global triple norm, taking into account also the interface interaction terms, as

$$|||w_h|||^2 = \sum_{i=1}^N \Big(|||w_h|||_i^2 + \|\delta(\varepsilon, \beta)[w_h]\|_{\partial\Omega_i}^2 \Big), \tag{2.4}$$

where $\delta(\varepsilon, \beta) = \frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} + \frac{1}{2}|\beta \cdot n|$. In what follows, we will also make use of the quantity $\delta^+(\varepsilon, \beta) = \frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} + \frac{1}{2}\beta \cdot n^+$. The explicit dependence of δ and δ^+ from ε and β will be omitted later on when there is no ambiguity of notation. For the continuity of the bilinear form we will also use the modified norm

$$\begin{aligned}]w_h[^2 = \sum_{i=1}^N \Big(&\|\sigma_1^{\frac{1}{2}} w_h\|_{\Omega_i}^2 + \|\beta\|_{L^\infty(\Omega)} \|\tilde{h}^{-\frac{1}{2}} w_h\|_{\Omega_i}^2 + \|\varepsilon^{\frac{1}{2}} \nabla w_h\|_{\Omega_i}^2 \\ &+ \|(\beta \cdot n)^{+\ \frac{1}{2}} w_h\|_{\partial\Omega_i \setminus \partial\Omega}^2 + J_i(w_h, w_h) \Big) + \|(\tilde{h}\varepsilon)^{\frac{1}{2}} \nabla w_h \cdot n\|_{\partial\Omega_i}^2 + \|\delta(\varepsilon, \beta) w_h\|_{\partial\Omega_i}^2. \end{aligned} \tag{2.5}$$

To prove convergence of the discrete solutions of formulation (2.1) to the exact solution of (1.1) we will first prove three preliminary lemmas giving Galerkin orthogonality, coercivity, and approximability. Existence of discrete solutions follows by the coercivity and convergence and is proved in Theorem 2.12.

We first recall a trace inequality and the standard inverse inequality that we will use repeatedly:

$$\|v\|_{0,\partial K}^2 \leq C\Big(h_K^{-1} \|v\|_K^2 + h_K \|v\|_{1,K}^2\Big) \quad \forall v \in H^1(K), \tag{2.6}$$

$$\|\nabla v\|_K \leq C_{inv} h_K^{-1} \|v\|_K. \tag{2.7}$$

For a proof of (2.6) we refer to [25, p. 26], and for a proof of (2.7) we refer to [9].

LEMMA 2.6 (Galerkin orthogonality). *Let $u \in \cup_{i=1}^N H^2(\Omega_i)$ be the exact solution of* (1.1) *and u_h the solution to* (2.1). *Then there holds*

$$A(u - u_h, v_h) + B(u - u_h, v_h) + J(u - u_h, v_h) = 0 \quad \forall v_h \in V_h.$$

Proof. By assumption we have that $[-\varepsilon \nabla u \cdot n + \beta \cdot n u] = [u] = 0$ in the sense of traces, and since $u \in \cup_{i=1}^N H^2(\Omega_i)$ there holds $J(u, v_h) = 0$. Therefore using the equality $[ab] = [a]\{b\} + \{a\}[b]$ and the fact that $\{\varepsilon\nabla u\cdot n\}_w - \beta\cdot n^+ u = \{\varepsilon\nabla u\cdot n - \beta\cdot nu\}$ we have

$$
\begin{aligned}
(2.8)\quad & A(u,v_h)+B(u,v_h)+J(u,v_h)\\
&= A(u,v_h) - \frac{1}{2}\sum_{i=1}^N \langle\{\varepsilon\nabla u\cdot n\}_w,[v_h]\rangle_{\partial\Omega_i} + \sum_{i=1}^N \left\langle \beta\cdot n^+ u,[v_h]\right\rangle_{\partial\Omega_i}\\
&= A(u,v_h) - \frac{1}{2}\sum_{i=1}^N \langle\{\varepsilon\nabla u\cdot n - \beta\cdot nu\},[v_h]\rangle_{\partial\Omega_i}\\
&= A(u,v_h) - \frac{1}{2}\sum_{i=1}^N \int_{\partial\Omega_i\setminus\partial\Omega}[(\varepsilon\nabla u\cdot n-\beta\cdot nu)v_h]\,\mathrm{d}s - \langle\varepsilon\nabla u\cdot n-\beta\cdot nu, v_h\rangle_{\partial\Omega}.
\end{aligned}
$$

By an integration by parts in each subdomain we obtain

$$
\begin{aligned}
A(u,v_h) &= \sum_{i=1}^N \{(\varepsilon\nabla u,\nabla v_h)_{\Omega_i} - (u,\beta\cdot\nabla v) + ((\sigma-\nabla\cdot\beta)u,v)_{\Omega_i}\}\\
&= \sum_{i=1}^N(-\varepsilon\Delta u + \beta\cdot\nabla u + \sigma u, v_h)_{\Omega_i} + \sum_{i=1}^N \langle\varepsilon\nabla u\cdot n-\beta\cdot nu, v_h\rangle_{\partial\Omega_i}\\
&= \sum_{i=1}^N (f,v_h)_{\Omega_i} + \frac{1}{2}\sum_{i=1}^N\int_{\partial\Omega_i\setminus\partial\Omega}[(\varepsilon\nabla u\cdot n-\beta\cdot nu)v_h]\,\mathrm{d}s + \langle\varepsilon\nabla u\cdot n-\beta\cdot nu,v_h\rangle_{\partial\Omega}.
\end{aligned}
$$

It then follows from (2.8) that

$$A(u,v_h)+B(u,v_h)+J(u,v_h) = (f,v_h);$$

combining this equality with (2.1) completes the proof. ◻

LEMMA 2.7 (coercivity). *For the formulation* (2.1) *there holds*

$$c|||z_h||| \le A(z_h,z_h)+B(z_h,z_h)+J(z_h,z_h) \quad \forall z_h\in V_h.$$

Proof. We essentially only need to show that the weakly imposed boundary and interface conditions do not destroy coercivity. We have

$$
\begin{aligned}
&A(z_h,z_h)+B(z_h,z_h) = \sum_{i=1}^N\Bigg(\int_{\Omega_i}(\sigma-\nabla\cdot\beta)\,z_h^2\,\mathrm{dx} + \|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i} - (z_h,\beta\cdot\nabla z_h)_{\Omega_i}\\
(2.9)\quad &\qquad + \left\langle\beta\cdot n^+ z_h,[z_h]\right\rangle_{\partial\Omega_i} - \langle\{\varepsilon\nabla z_h\cdot n\}_w,[z_h]\rangle_{\partial\Omega_i} + \left\langle\frac{\gamma_{bc}\varepsilon}{\tilde{h}}[z_h],[z_h]\right\rangle_{\partial\Omega_i}\Bigg).
\end{aligned}
$$

Consider the third term on the right-hand side of (2.9). Integration by parts yields

$$
\begin{aligned}
(2.10)\quad \sum_{i=1}^N(\beta\cdot\nabla z_h,z_h)_{\Omega_i} &= -\frac{1}{2}(\nabla\cdot\beta\; z_h,z_h)_\Omega + \sum_{i=1}^N\frac{1}{2}\langle\beta\cdot n\, z_h,z_h\rangle_{\partial\Omega_i}\\
&= -\frac{1}{2}(\nabla\cdot\beta\; z_h,z_h)_\Omega + \sum_{i=1}^N\frac{1}{4}\left\langle\beta\cdot n,\;[z_h^2]\right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \frac{1}{2}\langle\beta\cdot n\;z_h,\;z_h\rangle_{\partial\Omega}.
\end{aligned}
$$

Applying (2.10) to the third term of (2.9) and using the equality $a(a-b) = \frac{1}{2}(a^2 - b^2 + (a-b)^2)$ we get

$$\begin{aligned}\sum_{i=1}^{N}&\Big(-(z_h, \beta\cdot\nabla z_h)_{\Omega_i} + \left\langle \beta\cdot n^+ z_h, [z_h]\right\rangle_{\partial\Omega_i}\Big)\\ &= \sum_{i=1}^{N}\left(\frac{1}{2}(\nabla\cdot\beta\ z_h, z_h)_{\Omega_i} - \frac{1}{4}\left\langle \beta\cdot n, [z_h^2]\right\rangle_{\partial\Omega_i\setminus\partial\Omega}\right.\\ &\quad\left. + \frac{1}{2}\left\langle |\beta\cdot n|\ z_h,\ z_h\right\rangle_{\partial\Omega} + \frac{1}{2}\left\langle \beta\cdot n^+, [z_h^2]\right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \frac{1}{2}\left\langle \beta\cdot n^+[z_h], [z_h]\right\rangle_{\partial\Omega_i\setminus\partial\Omega}\right).\end{aligned}\tag{2.11}$$

By observing that $\sum_{i=1}^{N}\frac{1}{2}\left\langle \beta\cdot n^+, [z_h^2]\right\rangle_{\partial\Omega_i\setminus\partial\Omega} = \sum_{i=1}^{N}\frac{1}{4}\left\langle \beta\cdot n, [z_h^2]\right\rangle_{\partial\Omega_i\setminus\partial\Omega}$ we conclude that

$$\begin{aligned}\sum_{i=1}^{N}&\Big(-(z_h, \beta\cdot\nabla z_h)_{\Omega_i} + \left\langle \beta\cdot n^+ z_h, [z_h]\right\rangle_{\partial\Omega_i}\Big)\\ &= \sum_{i=1}^{N}\left(\frac{1}{2}(\nabla\cdot\beta\ z_h, z_h)_{\Omega_i} + \frac{1}{2}\left\langle |\beta\cdot n|\ z_h,\ z_h\right\rangle_{\partial\Omega} + \frac{1}{2}\left\langle \beta\cdot n^+[z_h], [z_h]\right\rangle_{\partial\Omega_i\setminus\partial\Omega}\right).\end{aligned}\tag{2.12}$$

We now consider the second, fifth, and sixth terms of (2.9). The nonsymmetric boundary integral is split using a Cauchy–Schwarz inequality followed by Young's inequality and controlled by the symmetric terms in the following fashion:

$$\begin{aligned}\sum_{i=1}^{N}&\left(\|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i} - \langle\{\varepsilon\nabla z_h\cdot n\}_w, [z_h]\rangle_{\partial\Omega_i} + \left\langle\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}}[z_h], [z_h]\right\rangle_{\partial\Omega_i}\right)\\ \geq \sum_{i=1}^{N}&\left(\|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i} - 2\alpha\|(\tilde{h}\varepsilon)^{\frac{1}{2}}\nabla z_h\cdot n\|^2_{\partial\Omega_i} + \left\langle\left(\gamma_{bc} - \frac{1}{4\alpha}\right)\frac{\{\varepsilon\}_w}{\tilde{h}}[z_h], [z_h]\right\rangle_{\partial\Omega_i}\right).\end{aligned}\tag{2.13}$$

As a consequence of the trace inequality (2.6) and inverse estimates we have

$$\|(\tilde{h}\varepsilon)^{\frac{1}{2}}\nabla z_h\cdot n\|^2_{\partial\Omega_i} \leq C_t\|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i},\tag{2.14}$$

and by choosing $\alpha = (4C_t)^{-1}$ and $\gamma_{bc} = 2C_t$ we conclude that

$$\begin{aligned}\sum_{i=1}^{N}&\left(\|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i} - \langle\{\varepsilon\nabla z_h\cdot n\}_w, [z_h]\rangle_{\partial\Omega_i} + \left\langle\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}}[z_h], [z_h]\right\rangle_{\partial\Omega_i}\right)\\ &\qquad\geq \frac{1}{2}\sum_{i=1}^{N}\left(\|\varepsilon^{\frac{1}{2}}\nabla z_h\|^2_{\Omega_i} + \left\langle\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}}[z_h], [z_h]\right\rangle_{\partial\Omega_i}\right).\end{aligned}\tag{2.15}$$

Combining the results of (2.9), (2.12), (2.15) and applying once again (2.14) and recalling the condition (1.2), the lemma follows, with a constant $c = \frac{1}{2}$. ∎

Remark 2.8. The constant C_t depends only on the mesh regularity and can be given an explicit expression in the case of piecewise linear elements (see [2]); for high order elements it can be computed by solving a small local eigenvalue problem (see [15]).

We will now proceed and prove approximability properties of the triple norm. The L^2-projection of u onto V_h will be denoted $\pi_h u$ and the nodal interpolation will be denoted $i_h u$. To avoid globally quasi-uniform meshes we need a stability estimate for the L^2-projection in weighted norms. This problem was considered in [13] and more recently in [3]. In [3] the following weighted stability estimate was proven:

$$\|\phi^* \pi_h u\|_\Omega \leq C\|\phi^* u\|_\Omega, \tag{2.16}$$

where ϕ^* is a piecewise linear weighting function satisfying

$$|\nabla \phi^*|_K| \leq \eta h_K^{-1} \max_{x\in K} \phi^* \tag{2.17}$$

for all K. Stability holds for η sufficiently small. We will use this stability result to prove the following

LEMMA 2.9. *If the polynomial order of the finite element space is k and $u \in H^{k+1}(\Omega)$, then there holds, for ρ sufficiently small,*

$$\sum_{K\in\mathcal{T}_{h,i}} (h_K^{-1}\|(\pi_h u - u)\|_K^2 + h_K\|\nabla(\pi_h u - u)\|_K^2) \leq C \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k+1}\|u\|_{k+1,K}^2. \tag{2.18}$$

Proof. First note that by adding and subtracting the nodal interpolant $i_h u$ in the H^1 contribution of (2.18) and applying a local inverse inequality we have

$$\begin{aligned} \sum_{K\in\mathcal{T}_{i,h}} & h_K\|\nabla(\pi_h u - u)\|_K^2 \\ &\leq C \sum_{K\in\mathcal{T}_{i,h}} (C_{inv}^2 h_K^{-1}\|(\pi_h u - i_h u)\|_K^2 + h_K\|\nabla(i_h u - u)\|_K^2). \end{aligned} \tag{2.19}$$

Hence it is sufficient to consider the L^2-part: $\sum_{K\in\mathcal{T}_{i,h}} h_K^{-\frac{1}{2}}\|(\pi_h u - u)\|_K^2$.

Take $\phi^* = \pi_h^* h_K^{-\frac{1}{2}}$. We must prove that this function satisfies (2.17) and that η can be made as small as needed by diminishing ρ. By the definition of the Oswald interpolant and the local quasi-regularity (1.4) one readily verifies that for all $K \in \mathcal{T}_{h,i}$

$$\max_{x\in K} |\nabla\phi^*| \leq h_K^{-1} \left| \max_{K'\in\mathcal{N}(K)} h_{K'}^{-\frac{1}{2}} - \min_{K'\in\mathcal{N}(K)} h_{K'}^{-\frac{1}{2}} \right| \leq h_K^{-1}(\rho^{\frac{1}{2}} - 1) \min_{K'\in\mathcal{N}(K)} (h_{K'}^{-\frac{1}{2}}).$$

Hence, using the inequality $\min_{K'\in\mathcal{N}(K)}(h_{K'}^{-\frac{1}{2}}) \leq \min_{x\in K}\phi^*$ we have $|\nabla\phi^*|_K| \leq (\rho^{\frac{1}{2}} - 1)h_K^{-1}\min_{x\in K}\phi^*$ on K, and therefore $\eta(\rho) = (\rho^{\frac{1}{2}} - 1)$ can be made arbitrarily small by choosing ρ small. Applying now the weighted stability estimate we have

$$\begin{aligned} \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\pi_h u - u)\|_K^2 &\leq \rho^{\frac{1}{2}}\|\phi^*(\pi_h u - u)\|_\Omega^2 \\ &\leq 2\rho^{\frac{1}{2}}(\|\phi^*(\pi_h u - i_h u)\|_\Omega^2 + \|\phi^*(i_h u - u)\|_\Omega^2) \\ &\leq C(\rho)\|\phi^*(i_h u - u)\|_\Omega^2 \leq C(\rho)\sum_K h_K^{2k+1}\|u\|_{k+1,\Omega}^2. \qquad \square \end{aligned}$$

LEMMA 2.10 (approximability). *Assume that the family of meshes $\mathcal{T}_{h,i}$ is locally quasi uniform with ρ such that Lemma 2.9 holds. Let $u \in \cup_{i=1}^N H^s(\Omega_i)$ with $s \geq k+1 \geq 2$ and let $\pi_h u$ denote the standard L_2-projection of u onto V_h; then we have that*

$$|||\pi_h u - u||| \leq C(\varepsilon^{\frac{1}{2}}\mathcal{H}(0,u) + \|\beta\|_{L^\infty(\Omega)}^{\frac{1}{2}}\mathcal{H}(1,u) + \sigma_0^{\frac{1}{2}}\mathcal{H}(2,u)),$$

where C is independent of σ, ε, β, and h but depends on the mesh geometry and

$$\mathcal{H}(\alpha, u) = \left(\sum_{i=1}^{N} \sum_{K \in \mathcal{T}_{h,i}} h_K^{2k+\alpha} \|u\|_{k+1,K}^2 \right)^{\frac{1}{2}}.$$

Proof. It follows from the stability of the L^2-projection and standard interpolation results that $\|\sigma_0^{\frac{1}{2}}(\pi_h u - u)\|_{\Omega_i} \leq \sigma_0^{\frac{1}{2}}(\sum_{K\in\mathcal{T}_{h,i}} h_K^{2(k+1)}\|u\|_{k+1,K}^2)^{\frac{1}{2}}$. We then write $\xi_h = \pi_h u - i_h u$, where i_h denotes the nodal interpolant, and note that $\xi_h = \pi_h(u - i_h u)$. By the H^1-stability of the L^2-projection on locally quasi-uniform meshes [4, 11] we may write

$$\|\nabla \xi_h\|_{\Omega_i} \leq \|\nabla(u - i_h u)\|_{\Omega_i} \leq C \left(\sum_{K\in\mathcal{T}_{h,i}} h_K^{2k} \|u\|_{k+1,K}^2 \right)^{\frac{1}{2}}. \tag{2.20}$$

It immediately follows by means of the triangular inequality that

$$\|\varepsilon^{\frac{1}{2}} \nabla(u - \pi_h u)\|_{\Omega_i}^2 \leq C\varepsilon \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k} \|u\|_{k+1,K}^2,$$

and by an application of the inverse inequality and Lemma 2.9 we have

$$\sum_{K\in\mathcal{T}_{h,i}} h_K^3 \|\nabla \xi_h\|_{1,\Omega}^2 \leq C \sum_{K\in\mathcal{T}_{h,i}} h_K \|\nabla \xi_h\|_{\Omega}^2 \leq \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k+1} \|u\|_{k+1,K}^2. \tag{2.21}$$

Using the trace inequality (2.6) together with (2.20) and (2.21), it follows that

$$\begin{aligned}
\|(\varepsilon \tilde{h})^{\frac{1}{2}} \nabla(\pi_h u - u)\cdot n\|_{\partial\Omega_i}^2 &\leq C \sum_{K\in\mathcal{T}_{h,i}} \left(\varepsilon \|\nabla(\pi_h u - u)\|_K^2 + \varepsilon h_K^2 \|\nabla(\pi_h u - u)\|_{1,K}^2 \right) \\
&\leq C\varepsilon \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k} \|u\|_{k+1,\Omega_i}^2.
\end{aligned}$$

Using once again (2.6), (2.20), and (2.21) we get in a similar fashion

$$\begin{aligned}
&J_1(u - \pi_h u, u - \pi_h u) \\
&\leq C \sum_{i=1}^{N} \gamma_{ip,i} \|\beta\|_{L^\infty(\Omega_i)} \sum_{K\in\mathcal{T}_{h,i}} \left(\|\tilde{h}^{\frac{1}{2}} \nabla(u - \pi_h u)\|_K^2 + \|\tilde{h}^{\frac{3}{2}} \nabla(u - \pi_h u)\|_{1,K}^2 \right) \\
&\leq \sum_{i=1}^{N} \|\beta\|_{L^\infty(\Omega_i)} \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k+1} \|u\|_{k+1,\Omega_i}^2.
\end{aligned}$$

Finally we note that for the boundary term we have, using (2.6) and (2.20),

$$\begin{aligned}
\langle \pi_h u - u, \pi_h u - u \rangle_{\partial\Omega_i} &\leq \sum_{K:\partial K\cap\partial\Omega_i \neq \emptyset} h_K^{-1} \|\pi_h u - u\|_K^2 + h_K \|\nabla(\pi_h u - u)\|_K^2 \\
&\leq \sum_{K\in\mathcal{T}_{h,i}} h_K^{2k+1} \|u\|_{k+1,\Omega_i}^2,
\end{aligned}$$

which concludes the proof. ∎

As an immediate consequence of the above result and Lemma 2.9 we have the following.

COROLLARY 2.11. *Under the same assumptions as in Lemma* 2.10 *we have that*

$$]\!]\pi_h u - u[\![\leq C(\varepsilon^{\frac{1}{2}}\mathcal{H}(0,u) + \|\beta\|^{\frac{1}{2}}_{L^\infty(\Omega)}\mathcal{H}(1,u) + \sigma_1^{\frac{1}{2}}\mathcal{H}(2,u)),$$

where C is independent of σ, ε, β, and h but depends on the mesh geometry.

THEOREM 2.12 (convergence). *Let $u \in \cup_{i=1}^N H^s(\Omega_i)$ with $s \geq k+1 \geq 2$ be the solution of* (1.1) *and let $u_h \in V_h$ be the solution of* (2.1). *Then the following a priori error estimate holds:*

$$|||u-u_h||| \leq C\left(\varepsilon^{\frac{1}{2}}\mathcal{H}(0,u) + \|\beta\|^{\frac{1}{2}}_{L^\infty(\Omega)}\mathcal{H}(1,u) + \left(\sigma_1^{\frac{1}{2}} + \frac{|\beta|_{W^{1,\infty}(\Omega)}}{\sigma_0}\right)\mathcal{H}(2,u)\right).$$

Proof. We decompose the error into two parts: $\eta = u - \pi_h u$ and $\xi_h = \pi_h u - u_h$. It follows that $u - u_h = \eta + \xi_h$. By Lemma 2.10 we know that

$$|||\eta||| \leq C(\varepsilon^{\frac{1}{2}}\mathcal{H}(0,u) + \|\beta\|^{\frac{1}{2}}_{L^\infty(\Omega)}\mathcal{H}(1,u) + \sigma_0^{\frac{1}{2}}\mathcal{H}(2,u)),$$

and it is therefore sufficient to study $\xi_h = \pi_h u - u_h$. Using Lemma 2.7 we have

$$c|||\xi_h|||^2 \leq A(\xi_h,\xi_h) + B(\xi_h,\xi_h) + J(\xi_h,\xi_h),$$

and by Galerkin orthogonality

$$c|||\xi_h|||^2 \leq A(\eta,\xi_h) + B(\eta,\xi_h) + J(\eta,\xi_h).$$

After an integration by parts in the convective term and an application of the Cauchy–Schwarz inequality in all other terms we have

$$c|||\xi_h|||^2 \leq]\!]\eta[\![\; |||\xi_h||| + |(\eta, \beta\cdot\nabla\xi_h)|.$$

Using now the orthogonality of the L^2-projection and Lemma 2.4 we may write

$$\begin{aligned} c|||\xi_h|||^2 &\leq]\!]\eta[\![\; |||\xi_h||| + |(\eta, \beta_h\cdot\nabla\xi_h - \pi^*\beta_h\cdot\nabla\xi_h)| + |(\eta,(\beta-\beta_h)\cdot\nabla\xi_h) \\ &\leq]\!]\eta[\![\; |||\xi_h||| + \|\beta\|_{L^\infty(\Omega)}\|\tilde{h}^{-\frac{1}{2}}\eta\| J(\xi_h,\xi_h)^{\frac{1}{2}} + |\beta|_{W^{1,\infty}(\Omega)}\|\eta\|\;\|\tilde{h}\nabla\xi_h\| \\ &\leq]\!]\eta[\![\; |||\xi_h||| + C_i\frac{|\beta|_{W^{1,\infty}(\Omega)}}{\sigma_0}\|\eta\|\; |||\xi_h|||. \end{aligned}$$

The theorem now follows by the approximation Lemma 2.10 and Corollary 2.11. ∎

Remark 2.13. The a priori error analysis carried out in this section holds true for any admissible choice of the weights w^+, w^- (such that $w^+, w^- > 0$ and $w^+ + w^- = 1$) that appear in the definition of $\{\cdot\}_w$ as also proved in [16] and [24]. In the following section we propose a definition of these weights according to the specific characteristics of the problem at hand.

2.2. Optimal choice of the averaging weights. To make the notation simpler, let us assume that only two subdomains Ω_i are considered with corresponding diffusivities ε_i, $i = 1, 2$. In this case, let $\partial\Omega_1 \setminus \partial\Omega$ be the interface between the subdomains and let n_1 be the outer normal with respect to Ω_1. Then we define the weighted average on the interface as $\{x(\xi)\}_w = \lim_{\delta\to 0}(w_1 x(\xi - n_1\delta) + w_2 x(\xi + n_1\delta))$.

The regularity assumptions on the solution u can be expected to hold only as long as $\varepsilon_i \geq \varepsilon_0 > 0$ in all the subdomains and the intersubdomain boundaries are

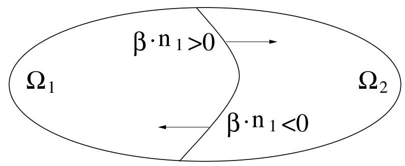

FIG. 2.1. *The model situation.*

smooth enough. In case ε_i vanishes in a subdomain, the weights w_i may be chosen so as to guarantee that the matching conditions automatically recover the physically correct behavior, relaxing the continuity of u but keeping the continuity of the fluxes. It turns out that balancing the diffusive fluxes yields a numerical scheme with the right asymptotic behavior if the diffusion coefficient vanishes in some subdomain. Let us exemplify this on a model case. We consider a domain Ω split into two neighboring subdomains Ω_1 and Ω_2 with a diffusion coefficient ε that is a regular function in each subdomain, but discontinuous across the interface $\partial\Omega_1 \cap \partial\Omega_2$. We choose the weights w_1 and w_2 such that

$$w_i(\xi) := \lim_{\delta\to 0} \frac{\varepsilon(\xi + \delta n_i)}{\varepsilon(\xi + \delta n_i) + \varepsilon(\xi - \delta n_i)} \quad \forall \xi \in \partial\Omega_1 \cap \partial\Omega_2,\ i = 1, 2, \tag{2.22}$$

where n_i is the outward unit normal with respect to Ω_i. We observe that such weights always satisfy $w_1(\xi) + w_2(\xi) = 1$ for all $\xi \in \partial\Omega_1 \cap \partial\Omega_2$. Moreover, in the case of smooth diffusivity across the interface, our choice coincides with the classical one, $w_1 = w_2 = \frac{1}{2}$. Furthermore, let us define $\omega(\xi) := w_1(\xi)\varepsilon_1(\xi) = w_2(\xi)\varepsilon_2(\xi)$. Our choice of the weights implies that $\{\varepsilon_i \nabla u_h \cdot n_i\}_w = 2\omega\{\nabla u_h \cdot n_i\}$, which shows that our method turns out to consider the arithmetic average of the gradients instead of the arithmetic average of the diffusion fluxes in order to construct the consistency term. Using these weights the coupling term between Ω_1 and Ω_2 becomes

$$\begin{aligned} B(u_h, v_h) = \sum_{i=1}^{2} \Bigg(& \left\langle \beta \cdot n^+ u_h, [v_h] \right\rangle_{\partial\Omega_i \setminus \partial\Omega} - \langle \omega\{\nabla u_h \cdot n\}, [v_h]\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ & - \langle \omega\{\nabla v_h \cdot n\}, [u_h]\rangle_{\partial\Omega_i\setminus\partial\Omega} + \left\langle \frac{\gamma_{bc} 2\omega}{\tilde{h}} [u_h], [v_h] \right\rangle_{\partial\Omega_i\setminus\partial\Omega} \Bigg). \end{aligned}$$

Consider the case when ε_1 goes to zero; then only the upwind flux term remains. One may readily verify that the coupling term $B(u_h, v_h)$ corresponds to the weak formulation of the conditions

$$\begin{aligned} -\varepsilon_2 \nabla u_{2,h} \cdot n_1 + \beta \cdot n_1 u_{2,h} = \beta \cdot n_1 u_{1,h} \text{ on } \ \partial\Omega_1 \setminus \partial\Omega, \text{ where } \beta \cdot n_1 > 0, \\ u_{1,h} = u_{2,h} \text{ and } -\varepsilon_2 \nabla u_{2,h} \cdot n_1 = 0 \text{ on } \partial\Omega_1 \setminus \partial\Omega, \text{ where } \beta \cdot n_1 < 0, \end{aligned}$$

which were proposed for the hybrid elliptic-hyperbolic coupling in Gastaldi and Quarteroni [14] (see also [10]). By the symmetry of the weights the same holds in the case $\varepsilon_2 = 0$. The convergence analysis of the iterative method and numerical experience also indicates that this choice of w_1 and w_2 is the only viable one for the iterative algorithm.

3. An iterative splitting method. To introduce and analyze the iterative method we will restrict the discussion to the case of two subdomains Ω_i, $i = 1, 2$, with interface $\partial\Omega_i \setminus \partial\Omega \neq \emptyset$. Nevertheless, the generalization to the multidomain case is straightforward and will be addressed later on. We denote with $u_{h,i} \in V_{h,k,i}$ the restriction on Ω_i of the global numerical solution. For the sake of simplicity, we also identify with $u_{h,i}$ the function on Ω that is obtained by extending $u_{h,i}$ to zero outside Ω_i. If we consider the formulation (2.1) and decouple the subdomains by using some approximation $u_{h,j}^k$ of $u_{h,j}$ with $j \neq i$ as boundary data from the neighboring subdomain with respect to Ω_i, we obtain the iterative scheme. Given $u_{h,1}^k, u_{h,2}^k$, for $k = 1, 2, \dots$, find $u_{h,1}^{k+1} \in V_{h,1}$ such that

$$A(u_{h,1}^{k+1}, v_{h,1}) + \tilde{B}(u_{h,1}^{k+1}, u_{h,2}^k, v_{h,1}) + J(u_{h,1}^{k+1}, v_{h,1}) + S(u_{h,1}^{k+1}, u_{h,1}^k, v_{h,1}) = (f, v_{h,1}) \tag{3.1}$$

and $u_{h,2}^{k+1} \in V_{h,2}$ such that

$$A(u_{h,2}^{k+1}, v_{h,2}) + \tilde{B}(u_{h,2}^{k+1}, u_{h,1}^k, v_{h,2}) + J(u_{h,2}^{k+1}, v_{h,2}) + S(u_{h,2}^{k+1}, u_{h,2}^k, v_{h,2}) = (f, v_{h,2}), \tag{3.2}$$

where

$$S(u_{h,i}^{k+1}, u_{h,i}^k, v_{h,i}) = \sum_{E \in G_h} \left\langle \frac{\gamma_{it}}{\tilde{h}} (u_{h,i}^{k+1} - u_{h,i}^k), v_{h,i} \right\rangle_E$$

are the terms that stabilize the iterations and the trace mesh is defined by

$$G_h = \{E \neq \emptyset : E = \partial K_i \cap \partial K_j;\ \forall K_i \in \mathcal{T}_{h,i};\ \forall K_j \in \mathcal{T}_{h,j};\ i \neq j\},$$

and we recall that $\tilde{h}(x)|_E = h_E$ for all $E \in G_h$.

The stabilization term $S(u_{h,i}^{k+1}, u_{h,i}^k, v_{h,i})$ corresponds to iteration relaxation and is mandatory to get good convergence properties. If S is omitted, we cannot prove convergence of the triple norm. In fact explicit control of the error in the jump over the interface is lost, and numerical experience shows very poor convergence as well for $S = 0$. Moreover, we note that the stabilization term is consistent in the sense that $S(u_{h,i}, u_{h,i}, v_{h,i}) = 0$. Finally, we have denoted with $\tilde{B}(u_{h,i}, u_{h,j}, v_{h,i})$, $i, j = 1, 2$, $j \neq i$, the interface/boundary penalty bilinear form after the iterative splitting, which is defined as follows:

$$\begin{aligned}
&\tilde{B}(u_{h,i}, u_{h,j}, v_{h,i}) = \langle \beta \cdot n_i^+ u_{h,i}, v_{h,i} \rangle_{\partial\Omega_i \setminus \partial\Omega} + \langle \beta \cdot n_i^- u_{h,j}, v_{h,i} \rangle_{\partial\Omega_i \setminus \partial\Omega} \\
&- \langle w_i \varepsilon_i \nabla u_{h,i} \cdot n_i + w_j \varepsilon_j \nabla u_{h,j} \cdot n_i, v_{h,i} \rangle_{\partial\Omega_i \setminus \partial\Omega} - \langle \varepsilon_i w_i \nabla v_{h,i} \cdot n_i, u_{h,i} - u_{h,j} \rangle_{\partial\Omega_i \setminus \partial\Omega} \\
&+ \left\langle 2 \frac{\gamma_{bc} \{\varepsilon\}_w}{\tilde{h}} (u_{h,i} - u_{h,j}), v_{h,i} \right\rangle_{\partial\Omega_i \setminus \partial\Omega} + \left\langle \frac{\gamma_{bc} \varepsilon}{\tilde{h}} u_{h,i}, v_{h,i} \right\rangle_{\partial\Omega_i \cap \partial\Omega} \\
&+ \langle \beta \cdot n^+ u_{h,i}, v_{h,i} \rangle_{\partial\Omega_i \cap \partial\Omega} - \langle \varepsilon_i \nabla u_{h,i} \cdot n, v_{h,i} \rangle_{\partial\Omega_i \cap \partial\Omega} - \langle \varepsilon_i \nabla v_{h,i} \cdot n, u_{h,i} \rangle_{\partial\Omega_i \cap \partial\Omega}.
\end{aligned}$$

Since the data on Ω_j are taken at the earlier iteration for both domains the two problems are decoupled and can be solved in parallel.

The present setting can easily be generalized to the case of several subdomains. Let $\bar{\Omega} = \cup_{i=1}^N \bar{\Omega}_i$ be the partition in N subdomains and let $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$ be the corresponding interfaces. Then, since the definition of A and J are already general with respect to N, problems (3.1) and (3.2) do not need to be modified in the multidomain case, provided that the definition of $\tilde{B}(u_{h,i}, u_{h,j}, v_{h,i})$ is adapted by replacing $\langle \cdot, \cdot \rangle_{\partial\Omega_i \setminus \partial\Omega}$ with $\sum_{i,j=1}^N \langle \cdot, \cdot \rangle_{\Gamma_{ij}}$. Thanks to the generality of the construction of G_h

the term $S(u_{h,i}, u_{h,i}, v_{h,i})$ remains unchanged. Moreover, in the multidomain case, the system of equations (3.1)–(3.2) should be complemented with one equation for each new subdomain. Although the formal generalization to the multidomain case is straightforward, we do not consider it here in order to reduce the notational complexity in the analysis of the iterative method.

LEMMA 3.1. *The subproblems* (3.1) *and* (3.2) *are well-posed in* $V_{h,i}$ *with respect to the norm* $|||\cdot|||_i$.

Proof. The proof is an immediate consequence of Lemma 2.7 restricted to one subdomain. ∎

We define the splitting error as $e_h^k = u_h - u_h^k$, where u_h is the solution to the finite element formulation of (2.1) and u_h^k is the solution after k iterations of (3.1) and (3.2). We will now state and prove the main result of this section.

THEOREM 3.2. *The iterative method defined by problems* (3.1) *and* (3.2) *converges when the relaxation parameter* γ_{it} *is chosen big enough. More precisely, there exists a positive constant* c (*that is, the coercivity constant of Theorem* 2.7) *such that*

(3.3)
$$c\sum_{k=1}^{\infty} |||e_h^k|||^2 \leq \sum_{i=1,2}\left(\frac{c}{2}\left\|\varepsilon_i^{\frac{1}{2}}\nabla e_{h,i}^0\right\|_{\Omega_i}^2 + \left\|(\delta^+)^{\frac{1}{2}} e_{h,i}^0\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 + \left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}} e_{h,i}^0\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\right)$$

Proof. By subtracting the decoupled formulation given by (3.1) and (3.2) from the formulation (2.1) we have

$$A(e_{h,1}^{k+1}, v_{h,1}) + \tilde{B}(e_{h,1}^{k+1}, e_{h,2}^k, v_{h,1}) + J(e_{h,1}^{k+1}, v_{h,1}) + S(e_{h,1}^{k+1}, e_{h,1}^k, v_{h,1}) = 0 \tag{3.4}$$

and

$$A(e_{h,2}^{k+1}, v_{h,2}) + \tilde{B}(e_{h,2}^{k+1}, e_{h,1}^k, v_{h,2}) + J(e_h^{k+1}, v_{h,2}) + S(e_{h,2}^{k+1}, e_{h,2}^k, v_{h,2}) = 0. \tag{3.5}$$

We now choose $v_{h,i} = e_{h,i}^{k+1}$ to obtain

$$\begin{aligned} A(e_h^{k+1}, e_h^{k+1}) + \tilde{B}(e_{h,1}^{k+1}, e_{h,2}^k, e_{h,1}^{k+1}) + \tilde{B}(e_{h,2}^{k+1}, e_{h,1}^k, e_{h,2}^{k+1}) + J(e_h^{k+1}, e_h^{k+1}) \\ + \sum_{i=1,2} S(e_{h,i}^{k+1}, e_{h,i}^k, e_{h,i}^{k+1}) = 0. \end{aligned}$$

Proceeding now by adding and subtracting $B(e_h^{k+1}, e_h^{k+1})$ we may write

$$\begin{aligned} A(e_h^{k+1}, e_h^{k+1}) + B(e_h^{k+1}, e_h^{k+1}) + J(e_h^{k+1}, e_h^{k+1}) + \sum_{i=1,2} S(e_{h,i}^{k+1}, e_{h,i}^k, e_{h,i}^{k+1}) \\ = B(e_h^{k+1}, e_h^{k+1}) - \tilde{B}(e_{h,1}^{k+1}, e_{h,2}^k, e_{h,1}^{k+1}) - \tilde{B}(e_{h,2}^{k+1}, e_{h,1}^k, e_{h,2}^{k+1}). \end{aligned} \tag{3.6}$$

The first three terms on the left-hand side will be controlled by the coercivity Lemma 2.7, while the term that stabilizes the iterations can be rewritten as follows:

$$\begin{aligned} \sum_{i=1,2} S(e_{h,i}^{k+1}, e_{h,i}^k, e_{h,i}^{k+1}) &= \sum_{i=1,2}\sum_{E\in G_h} \left\langle \frac{\gamma_{it}}{\tilde{h}}(e_{h,i}^{k+1} - e_{h,i}^k), e_{h,i}^{k+1}\right\rangle_E \\ &= \frac{1}{2}\sum_{i=1,2}\left[\left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}} e_{h,i}^{k+1}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 - \left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}} e_{h,i}^k\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 \right. \\ &\qquad \left. + \left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}} (e_{h,i}^{k+1} - e_{h,i}^k)\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\right]. \end{aligned} \tag{3.7}$$

It remains to bound the interface residual of the right-hand side:

$$R(e_{h,1}^k, e_{h,1}^{k+1}, e_{h,2}^k, e_{h,2}^{k+1}) = B(e_h^{k+1}, e_h^{k+1}) - \tilde{B}(e_{h,1}^{k+1}, e_{h,2}^k, e_{h,1}^{k+1}) - \tilde{B}(e_{h,2}^{k+1}, e_{h,1}^k, e_{h,2}^{k+1}).$$

The residual R is different from zero only on the interface of the subdomains and consists of three parts:

(A) the advective interface flux term from the advection term;
(B) the symmetric interface flux term from the Laplacian;
(C) the interface penalization term.

We now rearrange the terms for the three above-mentioned cases.

(A) The advective interface fluxes:

$$\begin{aligned}&\sum_{\substack{i,j=1,2\\ i\neq j}} \Big[\langle \beta\cdot n_i^+ e_{h,i}^{k+1}, e_{h,i}^{k+1} - e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle \beta\cdot n_i^+ e_{h,i}^{k+1}, e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ &\qquad\qquad - \langle \beta\cdot n_i^+ e_{h,i}^{k}, -e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big] \\ &= \sum_{\substack{i,j=1,2\\ i\neq j}} \Big[\langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,j}^{k+1} - e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} + \langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big].\end{aligned}$$

We observe that

$$\begin{aligned}&\langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,j}^{k+1} - e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ &\qquad\qquad \le \frac{1}{4\mu_i}\|(\beta\cdot n_i^+)^{\frac12}(e_{h,i}^k - e_{h,i}^{k+1})\|^2_{\partial\Omega_i\setminus\partial\Omega} + \mu_i\|(\beta\cdot n_i^+)^{\frac12}[e_h^{k+1}]\|^2_{\partial\Omega_i\setminus\partial\Omega}\end{aligned}$$

and

$$\begin{aligned}&\langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ &= \frac12\|(\beta\cdot n_i^+)^{\frac12} e_{h,i}^k\|^2_{\partial\Omega_i\setminus\partial\Omega} - \frac12\|(\beta\cdot n_i^+)^{\frac12} e_{h,i}^{k+1}\|^2_{\partial\Omega_i\setminus\partial\Omega} - \frac12\|(\beta\cdot n_i^+)^{\frac12}(e_{h,i}^{k+1} - e_{h,i}^k)\|^2_{\partial\Omega_i\setminus\partial\Omega}.\end{aligned}$$

By combining these results we obtain

$$\begin{aligned}&\sum_{\substack{i,j=1,2\\ i\neq j}} \Big[\langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,j}^{k+1} - e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} + \langle \beta\cdot n_i^+ (e_{h,i}^k - e_{h,i}^{k+1}), e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big] \\ &\quad \le \sum_{i=1,2}\Bigg[\mu_i\||\beta\cdot n_i|^{\frac12}[e_h^{k+1}]\|^2_{\partial\Omega_i\setminus\partial\Omega} + \frac{1-2\mu_i}{4\mu_i}\||\beta\cdot n_i|^{\frac12}(e_{h,i}^k - e_{h,i}^{k+1})\|^2_{\partial\Omega_i\setminus\partial\Omega} \\ &\qquad\qquad + \frac12\|(\beta\cdot n_i^+)^{\frac12} e_{h,i}^k\|^2_{\partial\Omega_i\setminus\partial\Omega} - \frac12\|(\beta\cdot n_i^+)^{\frac12} e_{h,i}^{k+1}\|^2_{\partial\Omega_i\setminus\partial\Omega}\Bigg].\end{aligned} \tag{3.8}$$

(B) The boundary part of the Laplacian operator may then be written as

$$\begin{aligned}&-\frac12\sum_{\substack{i,j=1,2\\ i\neq j}}\Big[2\langle\{\varepsilon\nabla e_h^{k+1}\cdot n_i\}_w, e_{h,i}^{k+1} - e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ &\quad -\langle\omega\nabla e_{h,i}^{k+1}\cdot n_i + \omega\nabla e_{h,j}^k\cdot n_i, e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle\omega\nabla e_{h,i}^{k+1}\cdot n_i, e_{h,i}^{k+1} - e_{h,j}^k\rangle_{\partial\Omega_i\setminus\partial\Omega} \\ &\quad -\langle\omega\nabla e_{h,i}^k\cdot n_i + \omega\nabla e_{h,j}^{k+1}\cdot n_i, -e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle\omega\nabla e_{h,j}^{k+1}\cdot n_i, e_{h,i}^k - e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big],\end{aligned}$$

which can be rewritten as follows:

$$
\begin{aligned}
-\frac{1}{2}\sum_{\substack{i,j=1,2\\ i\neq j}} \Big[&\langle \omega\nabla e_{h,j}^{k+1}\cdot n_i, e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle \omega\nabla e_{h,j}^{k}\cdot n_i, e_{h,i}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&- \langle \omega\nabla e_{h,i}^{k+1}\cdot n_i, e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} + \langle \omega\nabla e_{h,i}^{k}\cdot n_i, e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&+ \langle \omega\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i, e_{h,j}^{k+1} - e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&+ \langle \omega\nabla (e_{h,j}^{k+1} - e_{h,j}^{k})\cdot n_i, e_{h,i}^{k+1} - e_{h,i}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big],
\end{aligned}
$$

where we recall that $w_1\varepsilon_1 = w_2\varepsilon_2 = \omega$. For this choice of the averaging weights the first four terms vanish, precisely:

$$
\begin{aligned}
\sum_{\substack{i,j=1,2\\ i\neq j}} \Big[&\langle \omega\nabla e_{h,j}^{k+1}\cdot n_i, e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle \omega\nabla e_{h,j}^{k}\cdot n_i, e_{h,i}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&- \langle \omega\nabla e_{h,i}^{k+1}\cdot n_i, e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} + \langle \omega\nabla e_{h,i}^{k}\cdot n_i, e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big] = 0.
\end{aligned}
$$

By means of the Cauchy–Schwarz and Young inequalities, we have for the fifth term

$$
\begin{aligned}
\langle \omega\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i, e_{h,j}^{k+1} - e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega} &= \sum_{E\in G_h} \langle \omega^{\frac12}\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i, \omega^{\frac12}(e_{h,j}^{k+1} - e_{h,j}^{k})\rangle_E\\
&\leq \sum_{E\in G_h} 2\left[h_E^{\frac12}\|\omega^{\frac12}\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i\|_E \cdot h_E^{-\frac12}\|\omega^{\frac12}(e_{h,j}^{k+1} - e_{h,j}^{k})\|_E\right]\\
&\leq \sum_{E\in G_h} \left[\alpha_i h_E\|\omega^{\frac12}\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i\|_E^2 + (\alpha_i h_E)^{-1}\|\omega^{\frac12}(e_{h,j}^{k+1} - e_{h,j}^{k})\|_E^2\right].
\end{aligned}
$$

Then, by virtue of trace and inverse inequalities (see Remark 2.8), there exists a positive constant C_t such that

$$
\begin{aligned}
\sum_{E\in G_h} h_E\|\omega^{\frac12}\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i\|_E^2 &\leq C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\|\varepsilon_i^{\frac12}\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\|_{\Omega_i}^2\\
&\leq C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\left[\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k}\|_{\Omega_i}^2 + \|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k+1}\|_{\Omega_i}^2\right].
\end{aligned}
$$

We proceed analogously for the term $\langle \omega\nabla (e_{h,j}^{k+1} - e_{h,j}^{k})\cdot n_i, e_{h,i}^{k+1} - e_{h,i}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}$.

Summing up all the contributions we obtain that

$$
\begin{aligned}
(3.9)\quad -\frac{1}{2}\sum_{\substack{i,j=1,2\\ i\neq j}} \Big[&2\langle \omega\nabla (e_{h,i}^{k} - e_{h,i}^{k+1})\cdot n_i, e_{h,j}^{k+1} - e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&+ \langle \omega\nabla e_{h,j}^{k+1}\cdot n_i, e_{h,i}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} - \langle \omega\nabla e_{h,j}^{k}\cdot n_i, e_{h,i}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\\
&- \langle \omega\nabla e_{h,i}^{k+1}\cdot n_i, e_{h,j}^{k+1}\rangle_{\partial\Omega_i\setminus\partial\Omega} + \langle \omega\nabla e_{h,i}^{k}\cdot n_i, e_{h,j}^{k}\rangle_{\partial\Omega_i\setminus\partial\Omega}\Big]\\
\leq \sum_{i=1,2}\Bigg[&\alpha_i C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\left(\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k+1}\|_{\Omega_i}^2 + \|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k}\|_{\Omega_i}^2\right)\\
&+ \frac{\|w_i\varepsilon_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}}{\alpha_i}\|(\tilde{h})^{-\frac12}(e_{h,i}^{k+1} - e_{h,i}^{k})\|_{\partial\Omega_i\setminus\partial\Omega}\Bigg].
\end{aligned}
$$

(C) For the interface penalization term we get

$$\sum_{\substack{i,j=1,2\\ i\neq j}} \Bigg[\left\langle \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,j}^{k+1}), e_{h,i}^{k+1} - e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$- \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,j}^{k}), e_{h,i}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} - \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,i}^{k}), e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} \Bigg].$$

By means of algebraic manipulations we obtain

$$\left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,j}^{k}), e_{h,i}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,i}^{k}), e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$= \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,i}^{k}), e_{h,j}^{k+1} - e_{h,i}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,i}^{k}), e_{h,i}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$+ \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,j}^{k}), e_{h,i}^{k+1} - e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,j}^{k}), e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$+ \left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} [e_h^{k+1}] \right\|^2_{\partial\Omega_i\setminus\partial\Omega}.$$

By virtue of the particular choice of the weights that gives $2\omega = \{\varepsilon\}_w$ and by means of standard inequalities we observe that

$$\left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,i}^{k+1} - e_{h,i}^{k}), e_{h,j}^{k+1} - e_{h,i}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega} + \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,j}^{k}), e_{h,j}^{k+1} - e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$\leq \frac{1}{4\mu_i} \left[\left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} (e_{h,i}^{k+1} - e_{h,i}^{k}) \right\|^2_{\partial\Omega_i\setminus\partial\Omega} + \left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} (e_{h,j}^{k+1} - e_{h,j}^{k}) \right\|^2_{\partial\Omega_i\setminus\partial\Omega} \right]$$
$$\mu_i \left[\left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} [e_h^{k+1}] \right\|^2_{\partial\Omega_i\setminus\partial\Omega} + \left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} [e_h^{k+1}] \right\|^2_{\partial\Omega_i\setminus\partial\Omega} \right]$$
$$= \frac{1}{4\mu_i} \left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} (e_{h,i}^{k+1} - e_{h,i}^{k}) \right\|^2_{\partial\Omega_i\setminus\partial\Omega} + \mu_i \left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} [e_h^{k+1}] \right\|^2_{\partial\Omega_i\setminus\partial\Omega}$$

and that

$$\sum_{j=1,2} \left\langle \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right) (e_{h,j}^{k+1} - e_{h,j}^{k}), e_{h,j}^{k+1} \right\rangle_{\partial\Omega_i\setminus\partial\Omega}$$
$$= \frac{1}{2} \sum_{j=1,2} \Bigg[\left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} e_{h,j}^{k+1} \right\|^2_{\partial\Omega_i\setminus\partial\Omega} - \left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} e_{h,j}^{k} \right\|^2_{\partial\Omega_i\setminus\partial\Omega}$$
$$+ \left\| \left(\frac{\gamma_{bc}\omega}{\tilde{h}} \right)^{\frac{1}{2}} (e_{h,j}^{k+1} - e_{h,i}^{k}) \right\|^2_{\partial\Omega_i\setminus\partial\Omega} \Bigg]$$
$$= \frac{1}{2} \Bigg[\left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} e_{h,i}^{k+1} \right\|^2_{\partial\Omega_i\setminus\partial\Omega} - \left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} e_{h,i}^{k} \right\|^2_{\partial\Omega_i\setminus\partial\Omega}$$
$$+ \left\| \left(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde{h}} \right)^{\frac{1}{2}} (e_{h,i}^{k+1} - e_{h,i}^{k}) \right\|^2_{\partial\Omega_i\setminus\partial\Omega} \Bigg].$$

Summing up all the terms of the residual (C) we have

$$
(3.10)\quad \sum_{i=1,2}\Bigg[\Big\langle\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)(e_{h,i}^{k+1}-e_{h,j}^{k+1}),e_{h,i}^{k+1}-e_{h,j}^{k+1}\Big\rangle_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
-\Big\langle\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)(e_{h,i}^{k+1}-e_{h,j}^{k}),e_{h,i}^{k+1}\Big\rangle_{\partial\Omega_i\setminus\partial\Omega}-\Big\langle\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)(e_{h,j}^{k+1}-e_{h,i}^{k}),e_{h,j}^{k+1}\Big\rangle_{\partial\Omega_i\setminus\partial\Omega}\Bigg]
$$

$$
\leq \sum_{i=1,2}\Bigg[\frac{1-2\mu_i}{4\mu_i}\Big\|\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)^{\frac12}(e_{h,i}^{k+1}-e_{h,i}^{k})\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}+\mu_i\Big\|\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)^{\frac12}[e_h^{k+1}]\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+\frac12\Big\|\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)^{\frac12}e_{h,i}^{k}\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}-\frac12\Big\|\Big(\frac{\gamma_{bc}\{\varepsilon\}_w}{\tilde h}\Big)^{\frac12}e_{h,i}^{k+1}\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}\Bigg].
$$

By putting together (3.8), (3.9), (3.10) we obtain the following inequality:

$$
(3.11)\quad R(e_{h,1}^{k},e_{h,1}^{k+1},e_{h,2}^{k},e_{h,2}^{k+1})
$$

$$
\leq\sum_{i=1,2}\Bigg[\alpha_iC_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\Big(\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k+1}\|^2_{\Omega_i}+\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k}\|^2_{\Omega_i}\Big)+\mu_i\|\delta^{\frac12}[e_h^{k+1}]\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+\frac{1-2\mu_i}{4\mu_i}\|\delta^{\frac12}(e_{h,i}^{k+1}-e_{h,i}^{k})\|^2_{\partial\Omega_i\setminus\partial\Omega}+\frac{\|w_i\varepsilon_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}}{\alpha_i}\|\tilde h^{-\frac12}(e_{h,i}^{k+1}-e_{h,i}^{k})\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+\frac12\|(\delta^+)^{\frac12}e_{h,i}^{k}\|^2_{\partial\Omega_i\setminus\partial\Omega}-\frac12\|(\delta^+)^{\frac12}e_{h,i}^{k+1}\|^2_{\partial\Omega_i\setminus\partial\Omega}\Bigg].
$$

It should be noted that the right-hand side of (3.11) consists of terms that are either telescoping or of one of the following forms:

- terms containing $\nabla e_{h,i}^{k+1}$;
- terms containing a part $[e_h^{k+1}]$;
- terms containing a part $e_{h,i}^{k+1}-e_{h,i}^{k}$.

The first and second contributions will be controlled by the triple norm, and the last type of contributions will be controlled by the relaxation terms of (3.7). More precisely, by replacing (3.11) and (3.7) in (3.6) we obtain

$$
(3.12)\quad \sum_{i=1,2}\Bigg[c\|\sigma_0^{\frac12}e_{h,i}^{k+1}\|^2_{\Omega_i}+cJ(e_{h,i}^{k+1},e_{h,i}^{k+1})
$$

$$
+\Big(\frac{\gamma_{it}}{2}-\frac{\|w_i\varepsilon_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}}{\alpha_i}\Big)\|\tilde h^{-\frac12}(e_{h,i}^{k+1}-e_{h,i}^{k})\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
-\frac{1-2\mu_i}{4\mu_i}\|\delta^{\frac12}(e_{h,i}^{k+1}-e_{h,i}^{k})\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+(c-\alpha_iC_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)})\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k+1}\|^2_{\Omega_i}-\alpha_iC_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\|\varepsilon_i^{\frac12}\nabla e_{h,i}^{k}\|^2_{\Omega_i}
$$

$$
+c\|\delta^{\frac12}[e_h^{k+1}]\|^2_{\partial\Omega_i\cap\partial\Omega}+(c-\mu_i)\|\delta^{\frac12}[e_h^{k+1}]\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+\frac12\|(\delta^+)^{\frac12}e_{h,i}^{k+1}\|^2_{\partial\Omega_i\setminus\partial\Omega}-\frac12\|(\delta^+)^{\frac12}e_{h,i}^{k}\|^2_{\partial\Omega_i\setminus\partial\Omega}
$$

$$
+\frac12\Big\|\Big(\frac{\gamma_{it}}{\tilde h}\Big)^{\frac12}e_{h,i}^{k+1}\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}-\frac12\Big\|\Big(\frac{\gamma_{it}}{\tilde h}\Big)^{\frac12}e_{h,i}^{k}\Big\|^2_{\partial\Omega_i\setminus\partial\Omega}\Bigg]\leq 0.
$$

Then we choose the coefficients of Young's inequality, α_i and μ_i, as follows:

$$\alpha_i < \frac{c}{2C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}},\ \text{e.g., } \alpha_i = \frac{c}{4C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}};\quad \mu_i < c,\ \text{e.g., } \mu_i = \frac{c}{2};$$

and as a consequence of that, the relaxation parameter γ_{it} becomes

$$\begin{aligned}(3.13)\quad \gamma_{it} \geq{}& \frac{8C_t\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\|w_i\varepsilon_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}}{c}\\ &+\frac{\max[1-c,0]}{2c}\left(\gamma_{bc}\|\{\varepsilon\}_w\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)} + \|\beta\cdot n\tilde{h}\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\right),\ i=1,2.\end{aligned}$$

This allows us to rewrite (3.12) as follows:

$$\begin{aligned}\frac{c}{2}\left|\left|\left|e_h^{k+1}\right|\right|\right|^2 + \sum_{i=1,2}\Bigg[&\frac{c}{4}\left\|\varepsilon_i^{\frac{1}{2}}\nabla e_{h,i}^{k+1}\right\|_{\Omega_i}^2 - \frac{c}{4}\left\|\varepsilon_i^{\frac{1}{2}}\nabla e_{h,i}^{k}\right\|_{\Omega_i}^2\\ &+\frac{1}{2}\left\|(\delta^+)^{\frac{1}{2}}e_{h,i}^{k+1}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 - \frac{1}{2}\left\|(\delta^+)^{\frac{1}{2}}e_{h,i}^{k}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\\ &+\frac{1}{2}\left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}}e_{h,i}^{k+1}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 - \frac{1}{2}\left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}}e_{h,i}^{k}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\Bigg] \leq 0.\end{aligned}$$

Finally, summing up from $k=0$ to $k=M-1$, we obtain

$$\begin{aligned}c\sum_{k=0}^{M-1}\left|\left|\left|e_h^{k+1}\right|\right|\right|^2 &+ \sum_{i=1,2}\left(\left\|(\delta^+)^{\frac{1}{2}}e_{h,i}^{M}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 + \left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}}e_{h,i}^{M}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\right)\\ &\leq \sum_{i=1,2}\left(\frac{c}{2}\left\|\varepsilon_i^{\frac{1}{2}}\nabla e_{h,i}^{0}\right\|_{\Omega_i}^2 + \left\|(\delta^+)^{\frac{1}{2}}e_{h,i}^{0}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2 + \left\|\left(\frac{\gamma_{it}}{\tilde{h}}\right)^{\frac{1}{2}}e_{h,i}^{0}\right\|_{\partial\Omega_i\setminus\partial\Omega}^2\right),\end{aligned}$$

which implies (3.3). ∎

Remark 3.3. The general statement (3.13) implies the following choices of γ_{it}.

When $\varepsilon_i > 0$ for $i=1,2$ we have $c=\frac{1}{2}$ and $\gamma_{bc} \geq 2C_t$ in order to ensure coercivity. Then we insert $\gamma_{bc} = 2C_t$ into (3.13) and obtain

$$\gamma_{it} \geq 2C_t\left(1+8\|w_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}\right)\|w_i\varepsilon_i\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)} + \|\beta\cdot n\tilde{h}\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)},\ i=1,2.$$

For sufficiently small $\tilde{h}$ this expression can be summarized as $\gamma_{it} \simeq \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$. When $\varepsilon_1 = 0$ and $\varepsilon_2 > 0$ (or vice versa) we have $w_1 > 0$ and $w_2 = 0$. As a result of that the formula above becomes

$$\gamma_{it} \geq \frac{1}{2}\|\beta\cdot n\tilde{h}\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)},\ \text{i.e., } \frac{\gamma_{it}}{h} \geq \frac{1}{2}\|\beta\cdot n\|_{L^\infty(\partial\Omega_i\setminus\partial\Omega)}.$$

When $\varepsilon_1 = \varepsilon_2 = 0$ the coercivity constant becomes $c=1$. As a result of that (3.13) requires $\gamma_{it} \geq 0$.

4. Numerical results.

All the numerical experiments presented in this section were obtained using the `FreeFem++` library (http://www.freefem.org/ff++/index.htm).

4.1. Approximation and convergence properties of the iterative splitting method.

In this section we analyze the convergence of the iterative splitting method with respect to the mesh size $h = \max_{i=1,2}\max_{K\in\mathcal{T}_{h,i}} h_K$, the number of

TABLE 4.1
Convergence study with respect to h for conforming meshes.

Two subdomains, $h = 0.1$.

	P_1 FEM		P_2 FEM	
$\varepsilon = 1$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$
$2h$	$2.44\ 10^{-2}$	$5.82\ 10^{-1}$	$3.37\ 10^{-4}$	$5.05\ 10^{-2}$
h	$5.59\ 10^{-3}$	$2.65\ 10^{-1}$	4.62E-005	$1.26\ 10^{-2}$
Order	2.19	1.17	2.95	2.07
$\varepsilon = 10^{-3}$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$
$2h$	$1.65\ 10^{-2}$	$5.97\ 10^{-1}$	$8.88\ 10^{-4}$	$6.04\ 10^{-2}$
h	$3.64\ 10^{-3}$	$2.73\ 10^{-1}$	$1.02\ 10^{-4}$	$1.47\ 10^{-2}$
Order	2.24	1.16	3.21	2.10
$\varepsilon = 0$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$
$2h$	$1.69\ 10^{-2}$	$6.13\ 10^{-1}$	$9.95\ 10^{-4}$	$6.32\ 10^{-2}$
h	$3.80\ 10^{-3}$	$2.82\ 10^{-1}$	$1.23\ 10^{-4}$	$1.57\ 10^{-2}$
Order	2.22	1.16	3.10	2.07

Four subdomains $h = 0.08$.

	P_1 FEM		P_2 FEM	
$\varepsilon = 1$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$	$\lVert u-u_h\rVert_{0,\Omega}$	$\lVert u-u_h\rVert_{1,\Omega}$
$2h$	$1.50\ 10^{-2}$	$4.48\ 10^{-1}$	$1.76\ 10^{-4}$	$3.25\ 10^{-2}$
h	$3.38\ 10^{-3}$	$2.06\ 10^{-1}$	$1.65\ 10^{-5}$	$7.68\ 10^{-3}$
Order	2.15	1.12	3.42	2.08

subdomains N, and the value of the penalty parameters γ_{it}, γ_{bc}, and $\gamma_{ip,i}$, $i = 1, 2$, for different values of the diffusion parameters ε_i and of the transport field β. To this aim, we consider problem (1.1), where $\sigma = 1$ is fixed and f is chosen so that the exact solution is

$$u(x, y) = \exp(xy) \sin(\pi x) \sin(\pi y), \tag{4.1}$$

on a domain $\Omega =]0, 1[\times]0, 1[$ that has been split into $N = n^2$ subdomains such that $\bar{\Omega} = \cup_{i=1}^{N} \bar{\Omega}_i = \cup_{i_1,i_2=1}^{n} [(i_1 - 1)/n, i_1/n] \times [(i_2 - 1)/n, i_2/n]$, obtaining a checkerboard partition of size $H = 1/n$. The simplest case of two subregions $\bar{\Omega}_1 = [0, \frac{1}{2}] \times [0, 1]$ and $\bar{\Omega}_2 = [\frac{1}{2}, 1] \times [0, 1]$ is also addressed. For each subdomain, we introduce N quasi-uniform meshes $\mathcal{T}_{h,i}$ that can be either conforming or nonconforming on their interfaces, but for the tests presented here we consider conforming discretizations. For the comparison of different cases we choose $u^0_{h,i} = 0$ for $i = 1, \dots, N$ and consider a convergence test on the triple norm of the incremental error, namely, the iterations are stopped if $|||u_h^{k+1} - u_h^k|||/|||u_h^{k+1}||| \le tol$.

First of all, we aim to verify with numerical experiments the infinitesimal order with respect to h provided by Theorem 2.12. Table 4.1 shows that the optimal order of convergence is preserved for both linear and quadratic conforming elements. From now on, we will denote for simplicity $\|\cdot\|_{1,\Omega} \equiv (\sum_{i=1}^{N} \|\cdot\|_{1,\Omega_i})^{\frac{1}{2}}$.

Second, we aim to investigate the influence on the convergence rate of the iterative method of the parameters γ_{bc} and γ_{it} that appear in (3.1), (3.2). We study the number of iterations that the method needs to satisfy a tolerance $tol = 10^{-6}$ on the relative incremental error for several combinations of γ_{bc} and γ_{it}. Table 4.2 suggests

Table 4.2

The number of iterations necessary to converge with respect to a tolerance tol $= 10^{-6}$ *and on a quasi-uniform mesh of size* $h = 0.1$ *and a partition in two subdomains. Several combinations of the parameters* γ_{bc} *and* γ_{it} *in the case of the symmetric (right) and skewsymmetric coupling term (left) are addressed. In this case* $\varepsilon = 1$ *and* $\beta = [1, 1]$.

γ_{it}/γ_{bc}	$2\ 10^0$	$2\ 10^1$	$2\ 10^2$	$2\ 10^{-2}$	$2\ 10^{-1}$	$2\ 10^0$	$2\ 10^1$	$2\ 10^2$
$2\ 10^{-3}$	100	802	>1000	36	25	107	809	>1000
$2\ 10^{-2}$	101	803	>1000	34	25	107	810	>1000
$2\ 10^{-1}$	109	809	>1000	25	23	116	816	>1000
$2\ 10^0$	188	873	>1000	107	116	195	880	>1000
$2\ 10^1$	874	>1000	>1000	809	815	880	>1000	>1000
$2\ 10^2$	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000

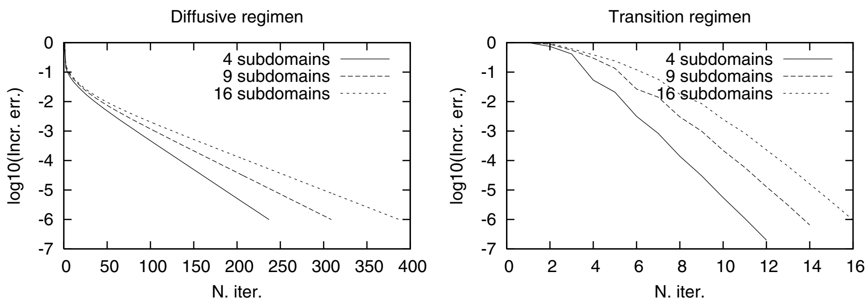

Fig. 4.1. *Convergence history of the iterative method for* $\varepsilon = 1.0$ *(right) and* $\varepsilon = 10^{-3}$ *(left) for the numerical tests on the coarse grids of Table* 4.3.

that an effective choice is to consider the small values of γ_{bc}, provided that the discrete problems (3.1), (3.2) remain well-posed according to Lemma 2.7. Recalling Remark 2.2, we analyze separately the symmetric and the nonsymmetric versions of the coupling term $B(u_h, v_h)$. In the symmetric case, Lemma 2.7 requires that $\gamma_{bc} = 2C_t$. In this case, Table 4.2 shows that the theoretical estimate obtained in Remark 3.3 is too restrictive for diffusion-dominated problems. Indeed, much smaller values of the estimated ones ensure better convergence properties. On the contrary, the numerical experiments presented in Table 4.5 suggest that the estimate of Remark 3.3 is effective for advection dominated problems. For the nonsymmetric case the limitations on γ_{bc} necessary for obtaining positivity of the discrete bilinear form change completely, in agreement with the analysis of interior penalty discontinuous Galerkin methods; see [1]. Indeed, only the restriction $\gamma_{bc} > 0$ is necessary. In this setting, the convergence properties of the iterative algorithm are much improved. Conversely, the approximation properties of the scheme are compromised since the discrete problem (2.1) is not adjoint consistent, and thus it does not enjoy optimal approximation properties in the L^2-norm (see [1] for a complete discussion). For the relaxation parameter γ_{it} we observe that in this case the choice $\gamma_{it} \simeq \gamma_{bc} \simeq 2\ 10^{-1}$ is effective.

The key point of this section is the characterization of the dependence of the convergence properties from the maximal mesh element size h and the number of subdomains N for different values of ε and β. More precisely, we analyze the diffusion dominated regimen ($\varepsilon = 1$), the transition regimen ($\varepsilon = 10^{-3}$), and the hyperbolic regimen ($\varepsilon = 0$). Indeed, Figure 4.1 and Table 4.3 show that the behavior of the method differs form one regimen to another. First of all, although Theorem 3.2 does

TABLE 4.3

The number of iterations necessary to converge with respect to a tolerance tol $= 10^{-6}$ *for several configurations of the partition in subdomains and several values of* ε. $\gamma_{bc} = 2$, $\gamma_{it} = \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$, *and* $\beta = [1, 1]$ *are fixed.*

H, N	1/2, 4		1/3, 9		1/4, 16	
h	0.13	0.06	0.12	0.06	0.12	0.07
$\varepsilon = 1$	237	445	309	579	388	723
Order h	−0.85		−0.91		−1.08	
Order H	–		−0.65		−0.77	
$\varepsilon = 10^{-3}$	12	14	14	17	16	20
Order h	−0.21		−0.28		−0.39	
Order H	–		−0.48		−0.56	
$\varepsilon = 0$	4	4	6	6	8	8
Order h	0		0		0	
Order H	–		−1		−1	

TABLE 4.4

The number of iterations necessary to converge with respect to a tolerance tol $= 10^{-6}$ *for different combinations of* β *and* ε *and for the case of* 16 *subdomains and* $h = 0.12$. $\gamma_{bc} = 2$, $\gamma_{it} = \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$ *are fixed.*

	$\varepsilon = 1$	$\varepsilon = 10^{-3}$	$\varepsilon = 0$
$\beta = [1, 1]$	388	16	8
$\beta = [1, 0]$	389	16	5
$\beta = [0, 1]$	389	16	5

not characterize the convergence behavior of the iterative method (3.1)–(3.2), Figure 4.1 puts into evidence that the incremental error is reduced according to the law C^k, where k is the iteration index and the constant $0 < C < 1$ is the convergence rate. Following this assumption, the number of iterations needed to satisfy a suitable tolerance on the incremental error is directly proportional to the convergence rate. As a consequence of that, Table 4.3 shows that in the diffusion-dominated regimen the convergence rate is inversely proportional to h and H. Following the heuristic motivations that are presented in [22] and [27], the inverse dependence on H can be explained observing that an iterative method that only exchanges information between neighboring subregions necessarily requires a number of steps to converge that is at least equal to the diameter of the dual graph corresponding to the subdomain partition, which is equivalent to $\mathcal{O}(H^{-1})$ when the diameter of Ω is unitary. The dual graph is constructed by introducing a vertex for each subregion and an edge between two subregions that share an interface. The inverse dependence on h is a consequence of (3.13) (see also Remark 3.3) which states that the relaxation term must be proportional to $\|\varepsilon\|_{L^\infty(\Omega)}/h$. Accordingly, by refining the mesh by a factor two, the number of iterations is doubled. Always in agreement with Remark 3.3 and with the fact that the relaxation term is allowed to vanish together with ε, the convergence rate of the method is less sensitive with respect to h for the transition case and completely insensitive with respect to the mesh size in the hyperbolic case. Indeed, when $\varepsilon = 0$ the number of iterations is only inversely dependent on H, and it is exactly equivalent to the number of steps that are needed to propagate the information along the diagonal of the checkerboard mesh defined by the subdomains, since the transport field is oriented along the diagonal. Furthermore, Table 4.4 suggests that these results do not deteriorate if the orientation of the transport field β is modified. Indeed, this is an advantage of the method proposed here with respect to the family of nonoverlapping domain decomposition methods arising from transmission conditions of Robin type, whose convergence may turn out to be slow when the transport field is tangential to the interface [21]. This benefit is due to the use of the upwind flux for the advection term. As a consequence of that, the corresponding transmission conditions are not symmetric with respect to β, in contrast to what happens for the family of methods inspired by transmission conditions of Robin type. Finally, we observe that in the hyperbolic case a multiplicative (Gauss–Seidel) iterative scheme is more preforming than the additive (Jacobi) method. For instance, since the subdomains in the checkerboard partition have been numbered by rows, when the transport field β is oriented in the vertical direction the multiplicative algorithm converges in 2 iterations, irrespectively of h and H.

4.2. Comparison of iterative methods. In order to assess the performance of the iterative method based on Nitsche's transmission conditions (denoted with a in Table 4.5 and defined by problems (3.1) and (3.2)) we compare it with the nonoverlapping Schwarz method proposed in [20] (denoted with b) and with the overlapping Schwarz method (denoted with c). For this comparison, we consider the test case proposed in the previous section where the domain Ω has been split into two subdomains, $\Omega_1 = [0, \frac{1}{2}] \times [0, 1]$ and $\Omega_2 = [\frac{1}{2}, 1] \times [0, 1]$. In the case of the overlapping Schwarz method we also introduce two overlapping domains, $\Omega_1^* = [0, \frac{1}{2} + \frac{1}{2}\delta] \times [0, 1]$, $\Omega_2^* = [\frac{1}{2} - \frac{1}{2}\delta, 1] \times [0, 1]$, and corresponding discretizations $\mathcal{T}^*_{h,i}$, $i = 1, 2$. Let $V^*_{h,i}$ be the finite element spaces defined on these meshes. Then, given $u^0_{h,i}$, for $k = 1, 2, \ldots$ we look for $u^k_{h,i} \in V^*_{h,i}$, $i = 1, 2$, such that

$$\begin{gathered}
A(u^{k+1}_{h,1}, v_{h,1}) + J(u^{k+1}_{h,1}, v_{h,1}) = (f_1, v_{h,1}) \;\; \forall v_{h,1} \in V^*_{h,1}, \quad \hat{u}^{k+1}_{h,1} = u^k_{h,2} \text{ on } \partial\Omega_1^* \cap \Omega_2^*, \\
A(u^{k+1}_{h,2}, v_{h,2}) + J(u^{k+1}_{h,2}, v_{h,2}) = (f_2, v_{h,2}) \;\; \forall v_{h,2} \in V^*_{h,2}, \quad \hat{u}^{k+1}_{h,2} = u^k_{h,1} \text{ on } \partial\Omega_2^* \cap \Omega_1^*, \\
u^{k+1}_{h,i} = \frac{1}{2}\hat{u}^{k+1}_{h,i} + \frac{1}{2}u^k_{h,i} \;\; i = 1, 2.
\end{gathered}$$

Recalling that the convergence of the overlapping Schwarz method can be accelerated by increasing the thickness of the overlapping region, that is, δ, we consider three cases, $\delta = \bar{h}$, $\delta = 2\bar{h}$, and $\delta = 4\bar{h}$, where $\bar{h}$ is the characteristic size of the quasi-uniform discretizations of Ω_1^* and Ω_2^*. The comparison with these cases will give a measure of the convergence performance of our method.

In Table 4.5, we compare the convergence and the approximation properties of these methods for the diffusion-dominated, the transition, and the hyperbolic regimens. The analysis of this table immediately shows that the method that we propose here is effective for the advection dominated and the hyperbolic regimens. In this case Nitsche's method a provides in general the best performances both for the convergence and the approximation properties for a fixed tolerance on the incremental error $tol = 10^{-6}$ and a given quasi-uniform mesh with $h = 0.05$.

In the diffusion-dominated case, the convergence of method a in the symmetric case is partially slowed down by the relaxation term. We have already observed that the choice $\gamma_{it} = \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$, motivated by the theoretical estimate derived in Remark 3.3, is not optimal. Indeed, the number of iterations needed to fulfill a tolerance of 10^{-6} on the incremental error is reduced from 354 to 190 if the parameter γ_{it} is divided by a factor of 100. In any case, this correction does not make method a with $s = 1$ (see Remark 2.2) competitive with method b in the diffusion-dominated

TABLE 4.5

The number of iterations necessary to converge with respect to a tolerance tol $= 10^{-6}$ *and the approximation error on a given quasi-uniform mesh characterized by* $h = 0.05$ *and a partition in* 2 *subdomains. Several instances of the iterative algorithms a, b, and c are considered. The instance of algorithm a with symmetric coupling terms is denoted with* $s = 1$*, while the nonsymmetric version is denoted with* $s = -1$*.*

Diffusion dominated regimen $\varepsilon = 1$, $\beta = [1, 1]$.

Method	N. iter.	$\|u - u_h\|_{0,\Omega}$	$\|u - u_h\|_{1,\Omega}$
a, $s = 1$, $\gamma_{bc} = 2$, $\gamma_{it} = \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$	354	$1.38\ 10^{-3}$	$1.32\ 10^{-1}$
a, $s = 1$, $\gamma_{bc} = 2$, $\gamma_{it} = 10^{-2}\gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$	190	$1.38\ 10^{-3}$	$1.32\ 10^{-1}$
a, $s = -1$, $\gamma_{bc} = 2\ 10^{-1}$, $\gamma_{it} = 2\ 10^{-1}$	43	$1.93\ 10^{-3}$	$1.25\ 10^{-1}$
a-hybrid	108	$1.37\ 10^{-3}$	$1.32\ 10^{-1}$
b	96	$1.37\ 10^{-3}$	$1.32\ 10^{-1}$
c, $\delta = \bar{h}$	210	$3.25\ 10^{-3}$	$1.27\ 10^{-1}$
c, $\delta = 2\bar{h}$	115	$2.35\ 10^{-3}$	$1.28\ 10^{-1}$
c, $\delta = 4\bar{h}$	65	$2.02\ 10^{-3}$	$1.36\ 10^{-1}$

Transition regimen $\varepsilon = 10^{-3}$, $\beta = [1, 1]$, $\gamma_{bc} = 2$, and $\gamma_{it} = \gamma_{bc}\|\varepsilon\|_{L^\infty(\Omega)}$.

Method	N. iter.	$\|u - u_h\|_{0,\Omega}$	$\|u - u_h\|_{1,\Omega}$
a, $s = 1$	12	$8.76\ 10^{-4}$	$1.33\ 10^{-1}$
a, $s = -1$	13	$8.75\ 10^{-4}$	$1.33\ 10^{-1}$
b	17	$1.03\ 10^{-3}$	$1.47\ 10^{-1}$
c, $\delta = \bar{h}$	46	$1.00\ 10^{-3}$	$1.37\ 10^{-1}$
c, $\delta = 2\bar{h}$	56	$1.27\ 10^{-3}$	$1.41\ 10^{-1}$
c, $\delta = 4\bar{h}$	42	$1.21\ 10^{-3}$	$1.51\ 10^{-1}$

Hyperbolic regimen $\varepsilon = 0$, $\beta = [1, 1]$, $\gamma_{bc} = 2$, and $\gamma_{it} = 0$.

Method	N. iter.	$\|u - u_h\|_{0,\Omega}$	$\|u - u_h\|_{1,\Omega}$
a, $s = \pm 1$	2	$9.48\ 10^{-4}$	$1.40\ 10^{-1}$
b	57	$2.44\ 10^{-3}$	$2.96\ 10^{-1}$
c, $\delta = \bar{h}$	52	$1.10\ 10^{-3}$	$1.45\ 10^{-1}$
c, $\delta = 2\bar{h}$	59	$1.48\ 10^{-3}$	$1.52\ 10^{-1}$
c, $\delta = 4\bar{h}$	45	$1.39\ 10^{-3}$	$1.63\ 10^{-1}$

case. Conversely, we observe that the convergence properties of the nonsymmetric version of method a is very satisfactory, while the approximation error in the L^2-norm reflects the suboptimality of this method. By comparing the properties of the symmetric and the nonsymmetric versions of method a, we observe that it may be possible to blend the benefits of the two methods by setting up a hybrid strategy (see Table 4.5, method a-hybrid). This consists in applying method a with $s = -1$, $\gamma_{bc} = \gamma_{it} = 2\ 10^{-1}$ until the tolerance equal to 10^{-6} is satisfied on the relative incremental error. As reported in Table 4.5, this procedure requires 43 iterations. Then, starting from the discrete solution computed in this way, we apply method a with $s = 1$, $\gamma_{bc} = 2$, $\gamma_{it} = 2\ 10^{-1}$ in order to improve the approximation error. This method requires 65 additional iterations to converge, and it reduces the L^2 approximation error of the nonsymmetric case from $1.93\ 10^{-3}$ to $1.37\ 10^{-3}$, which is equivalent to the error of the symmetric case. Since it is accurate and converges rapidly, the hybrid method outperforms both the symmetric and the nonsymmetric versions of method a. In the diffusive case, the hybrid method turns out to be almost equivalent to method b. These considerations promote further studies of the hybrid method and suggest investigating in detail whether the nonsymmetric formulation might be applied as a preconditioner for the symmetric case. Finally, a heuristic comparison with the overlapping Schwarz methods c suggests that method b behaves as an additive overlapping Schwarz algorithm with a relatively generous overlap of magnitude $\delta = 2\bar{h} \equiv 6\%$ of the diameter of Ω. On the other hand, the symmetric Nitsche's method a is almost equivalent to the overlapping method with small overlap $\delta = \bar{h} \equiv 3\%$.

From the point of view of computational cost we observe that the scheme (3.1)–(3.2) requires more effort for the construction of the finite element matrix corresponding to the coupling terms $B(u_h, v_h)$ than the family of Robin–Robin methods. Indeed, for the Robin–Robin methods the coupling matrix is easily constructed since it corresponds to a mass matrix on the degrees of freedom at the interface. Moreover in our case the bandwidth of the coupling matrix is increased because of the presence of first order derivatives in the coupling terms. This drawback is balanced by the fact that basic Robin–Robin iterative splitting methods preserve the optimal approximation properties of Lagrangian finite elements only if a superpenalty technique is applied; see [8]. This technique, however, compromises the convergence properties of the iterative algorithm.

4.3. Approximation of problems with discontinuous coefficients. In this section, we apply the numerical scheme (3.1)–(3.2) for the approximation of advection diffusion problems with discontinuous coefficients. To this purpose, the domain Ω has been split into two subdomains, $\Omega_1 = [0, \frac{1}{2}] \times [0, 1]$ and $\Omega_2 = [\frac{1}{2}, 1] \times [0, 1]$ with $\varepsilon(x) = 1.0$ for $x \in \Omega_1$ and $\varepsilon(x) = 2\ 10^{-2}$ for $x \in \Omega_2$. In the case $\sigma = 0$ and $f = 0$, the exact solution on each subregion Ω_1, Ω_2 can be easily expressed as an exponential function with respect to the x coordinate independently from the y coordinate. The global solution $u(x, y)$ is provided by choosing the value at the interface $x = \frac{1}{2}$ in order to ensure the following matching conditions:

$$\lim_{x \to \frac{1}{2}^-} u(x,y) = \lim_{x \to \frac{1}{2}^+} u(x,y) \quad \text{and} \quad \lim_{x \to \frac{1}{2}^-} -\varepsilon(x)\partial_x u(x,y) = \lim_{x \to \frac{1}{2}^+} -\varepsilon(x)\partial_x u(x,y).$$

More precisely, we set $u(0, y) = 1$, $u(1, y) = 0$, and by consequence of the matching conditions, we obtain

$$u\left(\frac{1}{2}, y\right) = \left[\frac{u(0,y)\exp(\frac{\beta}{2\varepsilon_1})}{1-\exp(\frac{\beta}{2\varepsilon_1})} + \frac{u(1,y)}{1-\exp(\frac{\beta}{2\varepsilon_2})}\right]\left[\frac{\exp(\frac{\beta}{2\varepsilon_1})}{1-\exp(\frac{\beta}{2\varepsilon_1})} + \frac{1}{1-\exp(\frac{\beta}{2\varepsilon_2})}\right]^{-1}.$$

As a result of that, the exact solution in each subdomain can be expressed as

$$u_1(x,y) = \frac{u(\frac{1}{2},y) - \exp(\frac{\beta}{2\varepsilon_1})u(0,y) + [u(0,y) - u(\frac{1}{2},y)]\exp(\frac{\beta x}{\varepsilon_1})}{1-\exp(\frac{\beta}{2\varepsilon_1})},$$

$$u_2(x,y) = \frac{u(1,y) - \exp(\frac{\beta}{2\varepsilon_2})u(\frac{1}{2},y) + [u(\frac{1}{2},y) - u(1,y)]\exp(\frac{\beta(x-\frac{1}{2})}{\varepsilon_2})}{1-\exp(\frac{\beta}{2\varepsilon_2})}.$$

The resulting function is represented in Figure 4.2. We aim to compare on the test problem defined above the accuracy of the scheme (3.1)–(3.2) with linear elements, precisely $V_h = \sum_{i=1}^{2} V_{h,1,i}$ (denoted by A) with the classical lagrangian linear elements over the whole domain Ω (denoted by B). We point out that in both cases the

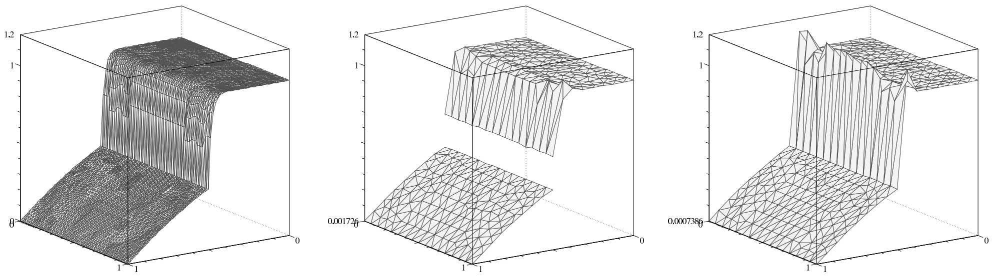

FIG. 4.2. *The nodal interpolant on a very refined mesh of the exact solution u of the test problem at hand* (*left*)*. The numerical approximation u_h obtained with method A* (*middle*) *and method B* (*right*) *in the case of the discretization characterized by $h_1 = 0.1$.*

TABLE 4.6
The quantitative comparison of the accuracy of methods A and B. The L^2-norm, $\|u_h - u\|_{0,\Omega}$, the H^1-norm, $\|u_h - u\|_{1,\Omega}$, and the maximum norm, $\|u_h - u\|_{L^\infty(\Omega)}$, are displayed.

	$\|u - u_h\|_{0,\Omega}$		$\|u - u_h\|_{1,\Omega}$		$\|u - u_h\|_{L^\infty(\Omega)}$	
h	Method A	Method B	Method A	Method B	Method A	Method B
0.1	1.81 10^{-2}	2.78 10^{-2}	2.07	1.78	2.13 10^{-1}	1.71 10^{-1}
0.05	6.98 10^{-3}	8.95 10^{-3}	1.31	1.06	1.07 10^{-1}	6.32 10^{-2}
0.026	2.56 10^{-3}	2.66 10^{-3}	7.49 10^{-1}	5.82 10^{-1}	1.40 10^{-1}	2.34 10^{-2}

continuous interior penalty stabilization method with $\gamma_{ip,i} = 2\ 10^{-2}$ has been applied to cure the instability of finite elements in the case of advection-dominated problems. We compare the two schemes on a family of quasi-uniform triangulations on Ω_1 and Ω_2 that are conforming at the interface of the subdomains and are characterized by a decreasing maximal element size $h_1 = 0.1$, $h_2 = 0.05$, and $h_3 = 0.026$. The quantitative analysis of the accuracy is based on the following indicators: the L^2-norm of the error, $\|u_h - u\|_{0,\Omega}$; the H^1-norm, $\|u_h - u\|_{1,\Omega}$, which is well defined since $u \in H^1(\Omega)$; and the maximum norm, $\|u_h - u\|_{L^\infty(\Omega)}$. The quantitative data are reported in Table 4.6, while a visual comparison is given in Figure 4.2. The analysis of the results suggests that the scheme (3.1)–(3.2) performs well for the approximation of problems with discontinuous coefficients when the mesh size is not small enough to fully resolve the boundary layers arising in the neighborhood of the region of discontinuity. The benefit of the scheme presented here with respect to the application of classical Lagrangian elements over Ω emerges if we consider the L^2-norm. For the mesh size h_1 method A provides numerical solutions that are smoother than method B (see Figure 4.2), where spurious oscillations appear in the neighborhood of the boundary layer that arise because of the discontinuity of ε. However, we observe that the L^∞ error of method B is smaller than in the case of method A, since for this method L^∞ errors arise when the very steep boundary layer across the discontinuity of ε is approximated with a jump. Finally, the analysis of the H^1-norm of the errors suggests that method B seems to be more prone to approximate the gradients of the solution in the boundary layer, although this benefit is effective when the computational mesh becomes fine enough to reasonably approximate the boundary layer.

5. Concluding remarks. In conclusion, the discretization scheme and the associated iterative method that we have proposed here turn out to be appealing for advection-dominated problems and in the case of discontinuous coefficients. Indeed, in these cases the method is competitive from the point of view of both computational effort and accuracy. A key role for the good properties when treating such problems is played by the average weights and the upwind treatment of the advection term in the interior penalty strategy applied for the coupling of the subdomains.

Acknowledgment. The authors are grateful for the referee's detailed and constructive criticisms and for the timely management of the manuscript.

REFERENCES

[1] 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.

[2] R. Becker, P. Hansbo, and R. Stenberg, *A finite element method for domain decomposition with non-matching grids*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 209–225.

[3] M. Boman, *Estimates for the L_2 projection onto continuous finite element spaces in a weighted L_p-norm*, BIT, 46 (2006), to appear.

[4] J. Bramble, J. Pasciak, and O. Steinbach, *On the stability of the L^2 projection in $H^1(\Omega)$*, Math. Comp., 71 (2002), pp. 147–156.

[5] E. Burman, *A unified analysis for conforming and nonconforming stabilized finite element methods using interior penalty*, SIAM J. Numer. Anal., 43 (2005), pp. 2012–2033.

[6] E. Burman, M. Fernández, and P. Hansbo, *Edge Stabilization: An Interior Penalty Method for the Incompressible Navier-Stokes Equation*, Tech. Rep. 23.2004, Ecole Polytechnique Federale de Lausanne, 2004.

[7] E. Burman and P. Hansbo, *Edge stabilization for Galerkin approximations of convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 1437–2453.

[8] T. Chacón-Rebollo and E. Chacón Vera, *Study of a non-overlapping domain decomposition method: Poisson and Stokes problems*, Appl. Numer. Math., 48 (2004), pp. 169–194.

[9] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, PA, 2002.

[10] J.-P. Croisille, A. Ern, T. Lelievre, and J. Proft, *Analysis and simulation of a coupled hyperbolic/parabolic model problem*, J. Numer. Math., 13 (2005), pp. 81–103.

[11] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp., 48 (1987), pp. 521–532.

[12] J. Douglas, Jr. and T. Dupont, *Interior penalty procedures for elliptic and parabolic Galerkin methods*, in Computing Methods in Applied Sciences (Second International Symposium, Versailles, 1975), Lecture Notes in Phys. 58, Springer-Verlag, Berlin, 1976, pp. 207–216.

[13] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems.* II. *Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[14] F. Gastaldi and A. Quarteroni, *On the coupling of hyperbolic and parabolic systems: Analytical and numerical approach*, Appl. Numer. Math., 6 (1989/90), pp. 3–31.

[15] P. Hansbo and M. G. Larson, *Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 1895–1908.

[16] B. Heinrich and K. Pietsch, *Nitsche type mortaring for some elliptic problem with corner singularities*, Computing, 68 (2002), pp. 217–238.

[17] R. Hoppe and B. Wohlmuth, *Element-oriented and edge-oriented local error estimators for nonconforming finite element methods*, RAIRO Modél. Math. Anal. Numér., 30 (1996), pp. 237–263.

[18] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.

[19] 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), pp. 1215–1238.

[20] G. Lube, L. Müller, and F. C. Otto, *A non-overlapping domain decomposition method for the advection-diffusion problem*, Computing, 64 (2000), pp. 49–68.

[21] F. Nataf and F. Rogier, *Factorization of the convection-diffusion operator and the Schwarz algorithm*, Math. Models Methods Appl. Sci., 5 (1995), pp. 67–93.

[22] A. Quarteroni and A. Valli, *Domain Decomposition Methods for Partial Differential Equations*, Numer. Math. Sci. Comput., Clarendon Press, Oxford University Press, New York, 1999.

[23] G. Rapin and G. Lube, *A stabilized three-field formulation for advection-diffusion equations*, Computing, 73 (2004), pp. 155–178.

[24] R. Stenberg, *Mortaring by a method of J. A. Nitsche*, in Computational Mechanics (Buenos Aires, 1998), CD-ROM, Centro Internac. Métodos Numér. Ing., Barcelona, 1998.

[25] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 1997.

[26] A. Toselli, *hp-finite element approximations on non-matching grids for partial differential equations with non-negative characteristic form*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 91–115.

[27] A. Toselli and O. Widlund, *Domain decomposition methods—algorithms and theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005.