# A MULTIGRID METHOD FOR UNFITTED FINITE ELEMENT DISCRETIZATIONS OF ELLIPTIC INTERFACE PROBLEMS*

THOMAS LUDESCHER†, SVEN GROSS†, AND ARNOLD REUSKEN†

**Abstract.** We consider discrete Poisson interface problems resulting from linear unfitted finite elements, also called cut finite elements. Three of these unfitted finite element methods, known from the literature, are studied. Two of these are suitable only for small jumps in the diffusion coefficient, and the third one has a robustness property that makes it appropriate also for interface problems with large coefficient jumps. All three methods rely on Nitsche's method to incorporate the interface conditions. The main topic of the paper is the development of a multigrid method, based on a novel prolongation operator for the unfitted finite element space and an interface smoother that is designed to yield robustness for large jumps in the diffusion coefficients. Numerical results are presented which illustrate efficiency of this multigrid method and demonstrate its robustness properties with respect to variation of the mesh size, location of the interface, and contrast in the diffusion coefficients.

**Key words.** geometric multigrid, unfitted finite elements, interface problem, Nitsche, CutFEM

**AMS subject classifications.** 65N30, 65F10, 65N22, 65N55

**DOI.** 10.1137/18M1203353

**1. Introduction.** In this paper we consider the elliptic interface problem

$$-\nabla \cdot (\mu \nabla u) = f \qquad \text{in } \Omega_i, \ i = 1, 2, \tag{1.1a}$$

$$u = u_D \qquad \text{on } \partial\Omega, \tag{1.1b}$$

$$[\![u]\!] = 0 \qquad \text{on } \Gamma, \tag{1.1c}$$

$$[\![-\mu\nabla u \cdot \mathbf{n}]\!] = 0 \qquad \text{on } \Gamma \tag{1.1d}$$

on a polygonal domain $\Omega = \Omega_1 \cup \Gamma \cup \Omega_2 \in \mathbb{R}^d$, $d = 2, 3$. The subdomains $\Omega_1$ and $\Omega_2$ are separated by the interface $\Gamma$, i.e., $\Gamma = \overline{\Omega}_1 \cap \overline{\Omega}_2$. For simplicity, we assume $\Gamma$ to be connected. The diffusion coefficient $\mu > 0$ is assumed to be piecewise constant, $\mu(x) = \mu_i = \text{const}$ on $\Omega_i$ with $\mu_1 \neq \mu_2$. The solution of problem (1.1) then exhibits a kink at the interface $\Gamma$. For a mesh which is *not* aligned to the interface, standard finite element methods can not approximate the solution $u$ accurately. Starting with the method introduced in [18], based on Nitsche's method to incorporate the interface conditions [28], several *un*fitted finite element methods, also called cut finite element methods (CutFEM), have been studied. These methods yield an optimal order of convergence for this class of elliptic interface problems. We refer to the recent overview paper [10] and the references therein for a detailed discussion of these methods.

Let $r_\mu := \frac{\max_{i=1,2} \mu_i}{\min_{i=1,2} \mu_i} \geq 1$ denote the ratio of the diffusion coefficient. In this paper we are interested in the efficient solution of problem (1.1) for two different regimes of $r_\mu$, namely, for the case of small jumps in the diffusion coefficient (i.e., $r_\mu \approx 1$) and for the case of large jumps in the diffusion coefficient (i.e., $r_\mu \gg 1$). For each of

†Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, D-52056 Aachen, Germany (ludescher@igpm.rwth-aachen.de, gross@igpm.rwth-aachen.de, reusken@igpm.rwth-aachen.de).

the cases we consider appropriate *un*fitted finite element discretizations and introduce and analyze an efficient multigrid method for the resulting discrete problems.

We restrict to piecewise linear finite elements. Higher order discretizations require an interface approximation of higher order, e.g., based on an isoparametric mapping [22, 24], which we do not treat here. Although several different unfitted finite element methods have been developed (cf. [10]), almost all of these methods use the same unfitted finite element space $V_\ell^\Gamma$, defined in (2.3). From the class of methods using this space we consider the following three important representatives. Firstly, as discretization method we use the original Nitsche unfitted finite element method [18], which was the starting point for the development of several variants. In this method one has to select an appropriate value for a stabilization parameter. It turns out that the performance of the multigrid method is very sensitive with respect to the choice of this parameter. For this reason we also consider a second unfitted finite element Nitsche type method introduced in [23], which is parameter-free. These first two methods are not robust with respect to large jumps in the diffusion coefficient $\mu$ and should only be applied to the case $r_\mu \approx 1$. Therefore, as a third unfitted finite element discretization we take the method from [11], which has a robustness property w.r.t. variation of the coefficient ratio $r_\mu$. This third method uses a so-called ghost penalty stabilization [9], which again requires the selection of a stabilization parameter value. The third method can also be applied to the case $r_\mu \approx 1$; the first two methods, however, are simpler to implement, which can be advantageous. We mention two further papers [3, 6] in which variants of these unfitted finite element methods for interface problems are studied. In these papers a particular choice of weights for the averaging operator (used in the Nitsche method) is introduced which leads to very strong robustness properties with respect to both large jumps in $\mu$ and the interface location. It turns out that for certain cut configurations and specific diffusion coefficient ratios the Nitsche penalty parameter becomes arbitrarily large (causing a corresponding condition number blow up for the stiffness matrix). This is not the case for the method from [9], and therefore we focus on that method.

Unfitted finite element methods are rather popular in fracture mechanics [15] (in that community these are often called extended finite element methods, XFEM). In [7, 16], algebraic multigrid methods for that problem class are developed. The performance of these methods for the interface problem (1.1c)–(1.1d) has not been studied.

It is well known (cf. [10]) that the stiffness matrices resulting from such unfitted finite element discretization methods may have a blow up in the condition number due to an instability caused by arbitrarily small cuts. For discretizations with higher order polynomials it is observed that the situation becomes even worse as cut basis functions can become linearly dependent within an element; see, e.g., [14]. This blow up effect can be avoided by a ghost penalty stabilization. As known from the literature, the methods from [18, 23] may suffer from this blow up effect, whereas (due to the ghost penalty stabilization) the method from [11] does not have this drawback. As we will see in this paper, the multigrid method that we propose does not deteriorate for the discretizations without ghost penalty stabilization, even though the condition number blows up due to small cuts.

There are only very few papers that we know of in which efficient iterative solvers for the discrete interface problems resulting from unfitted finite element methods are treated. In [25] a preconditioner for the original Nitsche unfitted finite element method [18] for (1.1) is developed which relies on a stable subspace splitting of the space $V_\ell^\Gamma$. This space is split into the standard linear finite element space and a complement

spanned by the cut basis functions. A block diagonal preconditioner is proposed that leads to a bounded condition number uniformly in the mesh size and the location of the interface. The condition number, however, deteriorates for larger jumps in the diffusion coefficient.

There are several papers in which multigrid methods are presented for discretizations of (1.1) that differ from the unfitted finite element methods considered in this paper. We mention a few of these. In [2, 1] a multigrid solver for the immersed interface method (IIM) [26] is presented, which uses nine point stencils on regular grids and a modified interpolation. A Krylov-accelerated multigrid method combined with the IIM is presented in [12]. In [27] a preconditioner for the IIM is developed that makes use of fast Poisson solvers. Recently, in [20] a geometric multigrid method has been studied for a finite element variant (on structured grids) of the IIM. For the special and simpler case, where the interface is aligned with the grid, a multigrid method with a specific coarsening strategy, such that the interface is aligned with coarse grids as well, is proposed in [30]. A multilevel preconditioner for a discontinuous Galerkin discretization of large contrast elliptic interface problems, with an interface aligned to the grid, is presented in [4].

There are a few papers in which fast solvers for unfitted finite element methods in the setting of fictitious domain techniques are treated, e.g., [5, 13].

To our knowledge an efficient geometric multigrid method for a class of unfitted finite element discretizations of (1.1) has not been treated in the literature. In this paper we present such a method. In this method there are two key ingredients that make it different from a standard multigrid method (for the Poisson equation). The first one is a novel prolongation operator that is tailor-made for the nonnested hierarchy of unfitted finite element spaces. Furthermore, for the case $r_\mu \gg 1$ we propose a smoothing procedure based on a combination of a standard Gauss–Seidel smoother with a specific local (close to the interface) correction procedure, which involves the inversion of a relatively low dimensional sparse matrix $A^\Gamma$. For this matrix we prove that (after diagonal scaling) its condition number is independent of $h$. This condition number, however, does deteriorate for large coefficient ratios $r_\mu$. Either a conjugate gradient iterative method (for moderate coefficient ratios) or a sparse direct solver (for large coefficient ratios) can be used for the efficient (approximate) solution of linear systems with the matrix $A^\Gamma$. The resulting multigrid solver, proposed in this paper, has arithmetic costs per iteration comparable to a few matrix-vector multiplications and a contraction number (much) smaller than one uniformly in the mesh size, the location of the interface, and the jump in the diffusion coefficient (where the latter property holds only for the robust discretization method from [11]). This optimality property is derived from numerical experiments for the three unfitted finite element discretizations outlined above.

An overview of the three unfitted discretization methods and the multigrid components for the different jumps in the diffusion coefficient is shown in Table 1.1.

The paper is organized as follows. In section 2 we introduce the three discretization methods from [18, 23, 11] that are considered. In section 3 we explain the multigrid method. In particular the new prolongation operator and a special (robust) smoother are introduced. The performance of the proposed multigrid method for the different unfitted finite element discretization methods is studied in section 4.

**2. Unfitted finite element discretizations.** In this section we describe three known unfitted finite element methods. These methods are the original Nitsche based CutFEM presented in [18, section 2.1], a variant without Nitsche stabilization

TABLE 1.1
*Overview of different discretization methods and multigrid components proposed for the cases of small and large jumps in the diffusion coefficient.*

|  | Small coefficient jump | Large coefficient jump |
|---|---|---|
| Discretization: | (Nitsche) from [18], (P-Nitsche) from [23] | (P-Nitsche) (for $r_\mu \leq 10^3$), ($\mu$-Nitsche) from [11] |
| Multigrid<br>• smoother: | Gauss–Seidel | Gauss–Seidel + interface correction (§3.2), $A^\Gamma$-solver:<br>• conjugate gradient solver (for $r_\mu \leq 10^2$)<br>• direct solver (for $r_\mu > 10^2$) |
| • prolongation: | tailor-made for (nonnested) spaces $V_\ell^\Gamma$ (§3.1) | |
| • restriction: | adjoint of prolongation | |
| • coarse grid op.: | discretization on coarser levels | |

parameter from [23, section 2.2], and a method from [11] which has better robustness properties w.r.t. the jump in the diffusion coefficient $\mu$ across the interface. In the first two methods we do *not* use any stabilization (e.g., ghost penalty) for improving the in general very poor conditioning of the system matrix. In the third method a ghost penalty stabilization is applied. *All three methods use the same unfitted finite element space $V_\ell^\Gamma$*, which is introduced below. We discuss certain elementary properties of the corresponding finite element isomorphism, which maps an element of $V_\ell^\Gamma$ to its representation in a particular basis of $V_\ell^\Gamma$. These properties are relevant for the derivation of the prolongation and restriction operators in the multigrid method in section 3.

Let $\{\mathcal{T}_\ell\}$, $\ell = 0, 1, \ldots$ be a family of nested, quasi-uniform simplicial triangulations of $\Omega$ which are not fitted to the interface $\Gamma$ ($\ell = 0$ corresponds to the coarsest level). The mesh size parameter corresponding to $\mathcal{T}_\ell$ is denoted by $h_\ell$. We introduce the standard space of continuous piecewise linear finite elements with zero boundary conditions

$$(2.1) \qquad V_\ell := \{\, v_h \in C(\Omega) \mid v_{h|\partial\Omega} = 0, \quad v_{h|T} \in \mathcal{P}_1 \quad \text{for all } T \in \mathcal{T}_\ell \,\}.$$

The unfitted variant of the space $V_\ell$ is obtained by "cutting finite element functions along $\Gamma$." In practice an approximation $\Gamma_\ell \approx \Gamma$ is used, which has a level dependent accuracy. In a multigrid solver one then has to deal with level dependent interface approximations $\Gamma_\ell$, where the approximation on a coarse level may be significantly different from the one on a fine level. This has to be taken into account in the prolongation and restriction operators used in the multigrid method. Actually, this is a key point in the particular multigrid method that we introduce in section 3. We assume approximations $\Gamma_\ell \approx \Gamma$, $\ell = 0, 1, 2, \ldots$, where $\Gamma_\ell$ is a connected interface approximation and define the set of cut elements $\mathcal{T}_\ell^\Gamma := \{T \in \mathcal{T}_\ell \mid \text{meas}_{d-1}(T \cap \Gamma_\ell) > 0\}$. We define a corresponding partitioning of the domain $\Omega = \Omega_{\ell,1} \cup \Gamma_\ell \cup \Omega_{\ell,2}$ with cut subdomains $\Omega_{\ell,i}$ and $\Gamma_\ell = \overline{\Omega}_{\ell,1} \cap \overline{\Omega}_{\ell,2}$; cf. Figure 2.1. Using the restriction operators to the subdomains $\Omega_{\ell,i}$

$$(2.2) \qquad \mathcal{R}_{\ell,i} v = \begin{cases} v_{|\Omega_{\ell,i}} & \text{on } \Omega_{\ell,i}, \\ 0 & \text{on } \Omega \setminus \Omega_{\ell,i}, \end{cases} \qquad i = 1, 2,$$
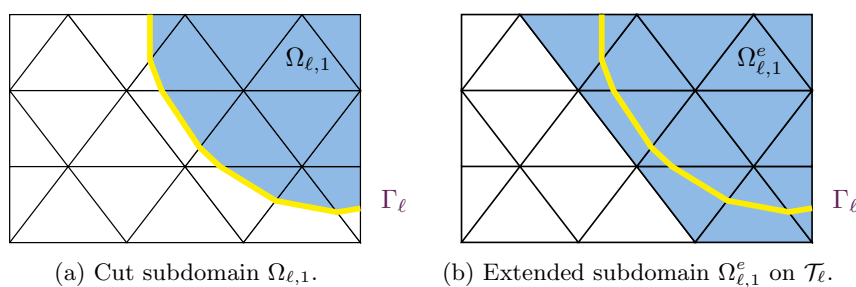
(a) Cut subdomain $\Omega_{\ell,1}$.      (b) Extended subdomain $\Omega_{\ell,1}^e$ on $\mathcal{T}_\ell$.

FIG. 2.1. *Cut and extended subdomains.*

we define the *unfitted finite element space*

$$(2.3) \qquad V_\ell^\Gamma = \mathcal{R}_{\ell,1} V_\ell \oplus \mathcal{R}_{\ell,2} V_\ell.$$

Note that in general, due to $\Gamma_{\ell-1} \neq \Gamma_\ell$, these spaces are *not nested*. For a description of the (nodal) basis in this space and the corresponding finite element isomorphism (mapping functions to the vector of coefficients in the basis) it is convenient to use an isomorphism between this unfitted finite element space $V_\ell^\Gamma$ and another finite element space; cf. (2.5) below. For this we first need some further notation. We introduce the *extended* domains $\Omega_{\ell,i}^e \supset \Omega_{\ell,i}$

$$\Omega_{\ell,i}^e := \cup \{ T \in \mathcal{T}_\ell \mid T \cap \Omega_{\ell,i} \neq \emptyset \}, \quad i = 1, 2,$$

(cf. Figure 2.1 for illustration).

We formulate assumptions on the interface approximations $\Gamma_\ell$ and the corresponding extended subdomains $\Omega_{\ell,i}^e$.

*Assumption* 1. For $\mathcal{T}_\ell$, $\Gamma_\ell$, $\ell = 0, 1, \ldots$ the following hold:
(A1) $\operatorname{dist}(\Gamma, \Gamma_\ell) \leq ch_\ell^2$;
(A2) $\Omega_{\ell,i}^e \subseteq \Omega_{\ell-1,i}^e, \quad i = 1, 2$;
(A3) for all $T \in \mathcal{T}_\ell$ with $T \cap \Gamma_\ell \neq \emptyset$ (hence, $T \cap \overline{\Omega}_{\ell,i} \neq \emptyset$ for $i = 1$ and 2) this intersection does not coincide with a subsimplex of $T$ (a face, edge, or vertex of $T$).

*Remark* 1. One particular choice for the approximate interface $\Gamma_\ell$ that is used in most applications of CutFEM is based on piecewise linear interpolation. If $\Gamma$ is characterized as the zero level of some level set function $\phi$, and $I_\ell \phi$ is the piecewise linear nodal interpolation in the finite element space $V_\ell$, one defines $\Gamma_\ell := \{ x \in \Omega \mid (I_\ell \phi)(x) = 0 \}$.

*Remark* 2. We briefly discuss Assumption 1. The assumption (A3), which is satisfied in generic cases, is made to simplify the presentation and avoid technical details. The geometric accuracy assumption (A1) is made to guarantee optimal second order discretization error bounds. The assumption (A2) is important for the construction of the multigrid prolongation operator. The assumption formalizes that the coarse level simplices in $\mathcal{T}_{\ell-1}$ that contain $\Gamma_\ell$ are the same as those containing $\Gamma_{\ell-1}$. Or in other words, if $\mathcal{T}_{\ell-1}^\Gamma$ denotes the subdomain formed by all simplices (on level $\ell - 1$) which are intersected by the interface approximation $\Gamma_{\ell-1}$, then the interface approximation $\Gamma_\ell$ on the next level remains in this subdomain. The assumptions (A1) and (A2) are satisfied for the example discussed in Remark 1.

Corresponding to the extended subdomains $\Omega_{\ell,i}^e$ we define the finite element spaces

$$(2.4) \qquad V_{\ell,i} := \{ (v_h)_{|\Omega_{\ell,i}^e} \mid v_h \in V_\ell \}, \quad i = 1, 2.$$

These are *standard* linear finite element spaces (on the triangulated *extended* subdomain $\Omega_{\ell,i}^e$). The standard nodal basis in $V_{\ell,i}$ is denoted by $(\varphi_{\ell,i}^k)_{1 \le k \le n_{\ell,i}}$, and the corresponding finite element isomorphism is given by

$$P_{\ell,i} : \mathbb{R}^{n_{\ell,i}} \to V_{\ell,i}, \quad P_{\ell,i}\mathbf{u}_i = \sum_{k=1}^{n_{\ell,i}} u_i^k \varphi_{\ell,i}^k.$$

One easily checks that

$$(2.5) \qquad \mathcal{R}_\ell : V_{\ell,1} \times V_{\ell,2} \to V_\ell^\Gamma, \quad \mathcal{R}_\ell \begin{pmatrix} u_{h,1} \\ u_{h,2} \end{pmatrix} := \mathcal{R}_{\ell,1} u_{h,1} + \mathcal{R}_{\ell,2} u_{h,2}$$

defines *an isomorphism.* As a first direct consequence of this we have that

$$(2.6) \qquad \{ \mathcal{R}_{\ell,i} \varphi_{\ell,i}^k \mid 1 \le k \le n_{\ell,i}, \ i = 1, 2 \}$$

is a basis of $V_\ell^\Gamma$. This basis is convenient for the derivation of a multigrid prolongation operator. Another useful basis of $V_\ell^\Gamma$ is introduced in section 3.2.

The natural finite element isomorphism $P_\ell^\Gamma : \mathbb{R}^{n_{\ell,1}+n_{\ell,2}} \to V_\ell^\Gamma$ is given by

$$(2.7a) \qquad P_\ell^\Gamma \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} := \mathcal{R}_\ell \begin{pmatrix} P_{\ell,1}\mathbf{u}_1 \\ P_{\ell,2}\mathbf{u}_2 \end{pmatrix}; \quad \text{hence}$$

$$(2.7b) \qquad (P_\ell^\Gamma)^{-1} \mathcal{R}_\ell \begin{pmatrix} u_{h,1} \\ u_{h,2} \end{pmatrix} = \begin{pmatrix} P_{\ell,1}^{-1} u_{h,1} \\ P_{\ell,2}^{-1} u_{h,2} \end{pmatrix}$$

holds. These relations will be useful for the construction of an appropriate multigrid prolongation operator. One easily checks that for $P_\ell^\Gamma$ defined in (2.7a) and $u_h \in V_\ell^\Gamma$, the vector $(P_\ell^\Gamma)^{-1} u_h$ contains the coefficients of the representation of $u_h$ in the nodal basis (2.6).

**2.1. Nitsche based CutFEM.** We briefly recall the method introduced in [18], which yields an accurate (optimal order) discretization for the case of small coefficient jumps, i.e., $r_\mu \approx 1$. Below $u_h \in V_\ell^\Gamma$, $(u_{h,1}, u_{h,2}) \in V_{\ell,1} \times V_{\ell,2}$ are always related by the isomorphism (2.5). We use the averaging operator (on level $\ell$)

$$\{\!\{u_h\}\!\} := (\kappa_1 u_{h,1} + \kappa_2 u_{h,2})|_{\Gamma_\ell}, \quad u_h \in V_\ell^\Gamma,$$

with $\kappa_i = \frac{|T_i|}{|T|}$, where $T_i := T \cap \Omega_{\ell,i}$ for $T \in \mathcal{T}_\ell$. The jump operator (on level $\ell$) is defined by

$$[\![u_h]\!] := (u_{h,1} - u_{h,2})|_{\Gamma_\ell}, \quad u_h \in V_\ell^\Gamma.$$

The Nitsche CutFEM for discretization of (1.1) is as follows: determine $u_h \in V_\ell^\Gamma$ such that

$$(\text{Nitsche}) \qquad a_h(u_h, v_h) = l(v_h) := (f, v_h)_\Omega \quad \text{for all} \ \ v_h \in V_\ell^\Gamma$$

with

$$(2.8) \qquad a_h(u,v) = a(u,v) + N^c(u,v) + N^c(v,u) + N^s_{\lambda_N}(u,v),$$

$$(2.9) \qquad a(u,v) := (\mu \nabla u, \nabla v)_{\Omega_{\ell,1} \cup \Omega_{\ell,2}},$$

$$(2.10) \qquad N^c(u,v) := (\{\!\!\{-\mu \nabla u \cdot \mathbf{n}_\ell\}\!\!\}, [\![v]\!])_{\Gamma_\ell},$$

$$(2.11) \qquad N^s_{\lambda_N}(u,v) := \frac{\lambda_N}{h_\ell}([\![u]\!], [\![v]\!])_{\Gamma_\ell}.$$

Note that the bilinear forms $a(\cdot,\cdot)$, $N^c(\cdot,\cdot)$, $N^s_{\lambda_N}(\cdot,\cdot)$ depend on the level $\ell$ through the level dependent approximation $\Gamma_\ell \approx \Gamma$. The vector $\mathbf{n}_\ell$ is the normal vector on $\Gamma_\ell$ (defined a.e.). Symmetry of the bilinear form in (Nitsche) is clear, and consistency, continuity, and coercivity are proved in [18]. The coercivity condition depends on the choice of the penalty parameter $\lambda_N$ which has to be taken "sufficiently large" (but independent of $\ell$). Together with an optimal approximation property of the space $V_\ell^\Gamma$ one obtains optimal order of convergence, i.e.,

$$(2.12) \qquad \|u^* - u_h\|_{0,\Omega} \le ch_\ell^2 \|u^*\|_{2,\Omega_1 \cup \Omega_2},$$

where $u^* \in H^2(\Omega_1 \cup \Omega_2)$ is the solution of problem (1.1). The constant $c$ in this error bound is independent of $\ell$ and the location of $\Gamma_\ell$ (but depends on $\mu$).

**2.2. Parameter-free CutFEM.** As mentioned in the previous section, the choice of a sufficiently large penalty parameter $\lambda_N$ is crucial for the coercivity of the bilinear form $a_h(\cdot,\cdot)$ used in the discretization (Nitsche). A drawback of the Nitsche method (Nitsche) is that it requires this manually chosen stabilization parameter. Furthermore, if $\lambda_N$ is chosen globally this may lead to overstabilization in some parts of the domain. In [23] a variant of the Nitsche method is presented that is parameter-free. We briefly describe this method and refer to [23] for more details and analysis of the method.

We use the bilinear forms defined in (2.9)–(2.11). The element bilinear forms, e.g., $a_T(\cdot,\cdot)$ corresponding to the element contribution to $a(\cdot,\cdot)$, are defined in the usual way: $a(\cdot,\cdot) = \sum_{T \in \mathcal{T}_\ell} a_T(\cdot,\cdot)$. On cut elements $T \in \mathcal{T}_\ell^\Gamma$, $T_i = T \cap \Omega_{\ell,i}$, a local lifting operator $\mathcal{L}_T : H^1(T_1 \cup T_2) \to \mathcal{P}_0^1(T_{1,2})$ is introduced, which maps into the space of polynomials of degree 1 that are orthogonal to constants on each subelement:

$$\mathcal{P}_0^1(T_{1,2}) := \{\, v \in L^2(T) \mid v|_{T_i} \in \mathcal{P}^1(T_i)/\mathcal{P}^0(T_i) \text{ for } i = 1,2 \,\}.$$

The lifting operator is constructed such that it has the property

$$(2.13) \qquad a_T(\mathcal{L}_T(u), v) = N_T^c(v,u) \quad \text{for all} \ \ v \in \mathcal{P}_0^1(T_{1,2}).$$

On uncut elements we set $\mathcal{L}_T(\cdot) = 0$ and thus obtain the global lifting operator

$$\mathcal{L} : H^1(\Omega_1 \cup \Omega_2) \to \bigoplus_{T \in \mathcal{T}_\ell^\Gamma} \mathcal{P}_0^1(T_{1,2}) \quad \text{with} \quad \mathcal{L}(u)|_T := \mathcal{L}_T(u).$$

The parameter-free bilinear CutFEM, which we call P-Nitsche, is as follows: determine $u_h \in V_\ell^\Gamma$ such that

$$\text{(P-Nitsche)} \qquad \tilde{a}_h(u_h, v_h) = l(v_h) := (f, v_h)_\Omega \quad \text{for all} \ \ v_h \in V_\ell^\Gamma$$

with

$$(2.14) \qquad \tilde{a}_h(u,v) = a(u,v) + N^c(u,v) + N^c(v,u) + 2a(\mathcal{L}(u), \mathcal{L}(v)) + N_1^s(u,v).$$

Through the lifting defined in (2.13) the choice $\lambda_N = 1$, as indicated by $N_1^s(\cdot, \cdot)$ in (2.14), suffices to guarantee coercivity. Sufficient additional stability is implicitly added by the term $a(\mathcal{L}(\cdot), \mathcal{L}(\cdot))$. A proof of consistency, coercivity, and continuity can be found in [23]. These properties imply the same error bound as in (2.12).

*Remark* 3. In assembling the stiffness matrix corresponding to the bilinear form (2.14), for the evaluation of $\mathcal{L}_T$ one has to solve the local systems in (2.13), which requires the inversion of element matrices. Further implementational aspects are discussed in [21, 23].

**2.3. Coefficient-stable CutFEM.** The methods introduced in section 2.1 and 2.2 are not robust with respect to large jumps in the diffusion coefficients, i.e., for $r_\mu = \mu_{\max}/\mu_{\min} \gg 1$. Hence, we also consider the method from [11] which has better robustness properties w.r.t. the variation of the coefficient jump $r_\mu$. In that method the averaging operator is defined as

$$\{\!\{u_h\}\!\}_\mu := (\kappa_1^\mu u_{h,1} + \kappa_2^\mu u_{h,2})|_{\Gamma_\ell}, \quad u_h \in V_\ell^\Gamma,$$

with harmonic weights

$$\kappa_1^\mu = \frac{\mu_2}{\mu_1 + \mu_2}, \quad \kappa_2^\mu = \frac{\mu_1}{\mu_1 + \mu_2} \quad (\text{hence, } \kappa_1^\mu + \kappa_2^\mu = 1).$$

We emphasize the harmonic averaging by the subscript $\mu$ in the averaging operator. It turns out that using these weights the method is not robust with respect to the interface location, i.e., there occur instabilities due to "small cuts." To recover the stability, additional terms are included in the bilinear form, which are based on the so-called ghost penalty stabilization [9]:

$$(2.15) \qquad g_h(u_h, v_h) := \sum_{i=1}^{2} \sum_{F \in \mathcal{F}_i^\ell} \varepsilon_g\, \mu_i\, h_F([\![\nabla u_{h,i} \cdot \mathbf{n}_F]\!]_F, [\![\nabla v_{h,i} \cdot \mathbf{n}_F]\!]_F)_F$$

with a (constant) stabilization parameter $\varepsilon_g > 0$. Here, $[\![\nabla v_{h,i} \cdot \mathbf{n}_F]\!]_F$ denotes the jump of the normal component of the piecewise constant function $\nabla v_{h,i}$ across the face $F$. The sets of faces only contain faces close to the interface and are defined as

$$\mathcal{F}_i^\ell := \{F \subset \partial T \mid T \in \mathcal{T}_\ell^\Gamma, F \not\subset \partial \Omega_{\ell,i}^e\} \quad \text{for } i = 1, 2.$$

An illustration for different triangulation levels is shown in Figure 2.2.

The resulting discretization method, which we call $\mu$-Nitsche, is as follows: determine $u_h \in V_\ell^\Gamma$ such that

$(\mu\text{-Nitsche})$ $\qquad\qquad \hat{a}_h(u_h, v_h) = l(v_h) := (f, v_h)_\Omega \quad \text{for all } v_h \in V_\ell^\Gamma$

with

$$\hat{a}_h(u, v) := a(u, v) + N^{c,\mu}(u, v) + N^{c,\mu}(v, u) + N_{\lambda^\mu}^s(u, v) + g_h(u, v),$$
$$N^{c,\mu}(u, v) := (\{\!\{-\mu\nabla u \cdot \mathbf{n}_\ell\}\!\}_\mu, [\![v]\!])_{\Gamma_\ell},$$
$$\lambda^\mu := \frac{2\mu_1\mu_2}{\mu_1 + \mu_2}\lambda_N.$$

(a) Face set $\mathcal{F}_1^{\ell-1}$ (in red) on $\mathcal{T}_{\ell-1}$          (b) Face set $\mathcal{F}_1^{\ell}$ (in red) on $\mathcal{T}_\ell$
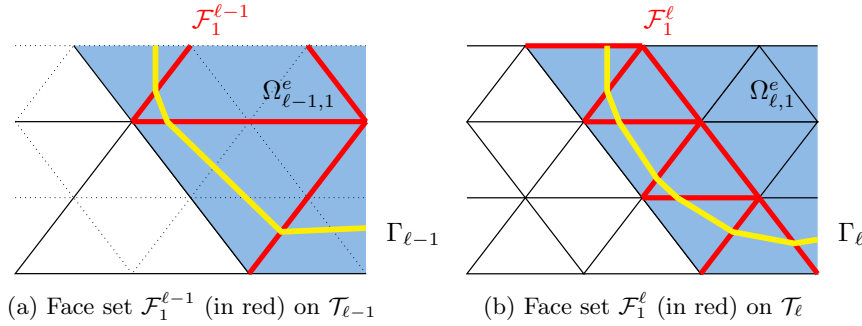
FIG. 2.2. *Face sets for ghost penalty stabilization on different refinement levels.*

The penalty parameter $\lambda^\mu$ is $\mu$-dependent (with $\lambda_N > 0$ an $\ell$- and $\mu$-independent parameter value). In [11] the estimate $\|\!|u^* - u_h|\!\|_{1,h,\Omega} \leq ch\|u^*\|_{H^2(\Omega_1\cup\Omega_2)}$ was proved with norm definition

$$\|\!|v_h|\!\|_{1,h,\Omega}^2 := \sum_{i=1}^{2} \|\mu_i^{\frac{1}{2}}\nabla v_{h,i}\|_{0,\Omega}^2 + \|\{\!\{\mu\}\!\}_\mu^{\frac{1}{2}}\{\!\{\nabla v_h\cdot\mathbf{n}_\Gamma\}\!\}\|_{-\frac{1}{2},h,\Gamma}^2 + \|\{\!\{\mu\}\!\}_\mu^{\frac{1}{2}}[\![v_h]\!]\|_{\frac{1}{2},h,\Gamma}^2,$$

where $\|v_h\|_{\pm\frac{1}{2},h,\Gamma}^2 := \sum_{T\in\mathcal{T}_\ell^\Gamma} h_T^{\mp 1}\|v_h\|_{0,\Gamma_T}^2$. This error bound can be combined with a duality argument, resulting in an optimal error estimate as in (2.12).

**3. Multigrid method.** In this section we present an efficient multigrid solver for the discrete problems resulting from (Nitsche), (P-Nitsche), or ($\mu$-Nitsche). These discretizations use the same unfitted finite element space $V_\ell^\Gamma$.

For a multigrid solver we have to specify how restrictions and prolongations are constructed and which smoother is used. We first address the restrictions. Given an appropriate prolongation (explained below) $\mathbf{p}_\ell$ (in matrix representation), we make the canonical choice for the restriction, namely, $\mathbf{r}_\ell := \mathbf{p}_\ell^T$. The coarse grid matrices are constructed by the usual procedure in geometric multigrid methods, namely, *direct discretization on the coarser levels*. Alternatively one could consider the construction of coarse grid matrices by the Galerkin approach ($A_{\ell-1} = \mathbf{r}_\ell A_\ell \mathbf{p}_\ell$). It turns out that both methods yield similar results; cf. Remark 8. In the following two subsections we discuss the prolongation operator and the smoother.

**3.1. Prolongation operator.** For the derivation and representation of the prolongation operator we use the basis (2.6). Recall that the unfitted finite element spaces are not necessarily nested. We introduce a natural prolongation $p_\ell : V_{\ell-1}^\Gamma \to V_\ell^\Gamma$. Note that due to (A2) the spaces $V_{\ell,i}$ are nested in the sense that for $w \in V_{\ell-1,i}$ we have $w_{|\Omega_{\ell,i}^e} \in V_{\ell,i}$, $i = 1, 2$. We introduce the operator

$$(3.1) \qquad I_\ell^\Gamma : V_{\ell-1,1}\times V_{\ell-1,2} \to V_{\ell,1}\times V_{\ell,2}, \quad I_\ell^\Gamma\begin{pmatrix}u_{H,1}\\u_{H,2}\end{pmatrix} := \begin{pmatrix}I_{\ell,1}^\Gamma u_{H,1}\\I_{\ell,2}^\Gamma u_{H,2}\end{pmatrix}$$

with

$$(3.2) \qquad I_{\ell,i}^\Gamma : V_{\ell-1,i}\to V_{\ell,i}, \quad I_{\ell,i}^\Gamma u_{H,i} := u_{H,i}|_{\Omega_{\ell,i}^e},$$

the natural embedding (by restriction) of a coarse level function $u_{H,i}$ (defined on $\Omega_{\ell-1,i}^e$) into $V_{\ell,i}$. Hence for $u_H = \mathcal{R}_{\ell-1}(u_{H,1}, u_{H,2})^T \in V_{\ell-1}^\Gamma$ the prolongation

$$(3.3) \qquad p_\ell u_H := \mathcal{R}_\ell I_\ell^\Gamma \mathcal{R}_{\ell-1}^{-1} u_H = \mathcal{R}_\ell\big(u_{H,1|\Omega_{\ell,1}^e}, u_{H,2|\Omega_{\ell,2}^e}\big)^T \in V_\ell^\Gamma$$

is well-defined. The prolongation procedure is illustrated in Figure 3.1.

(a) $\mathcal{R}_{\ell-1}^{-1} u_H$ on $\mathcal{T}_{\ell-1}$ (b) $\mathcal{R}_\ell^{-1} p_\ell u_H$ on $\mathcal{T}_\ell$
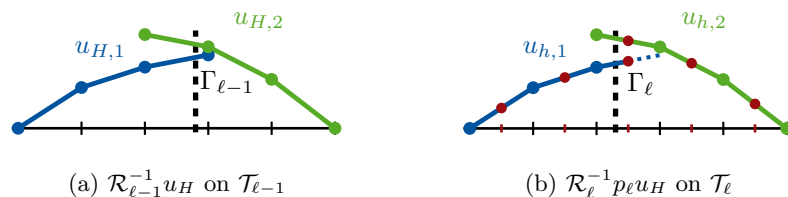
FIG. 3.1. *Illustration of prolongation procedure.*

Using the results in (2.7) we derive the matrix representation of this prolongation operator. For this we use the matrix representation of the prolongation corresponding to $V_{\ell,i}$, i.e.,

$$\mathbf{p}_{\ell,i} := P_{\ell,i}^{-1} I_{\ell,i}^\Gamma P_{\ell-1,i}.$$

Note that for this operator to be well-defined we need (A2), because for $\mathbf{u} \in \mathbb{R}^{n_{\ell-1,i}}$ we have that $w := P_{\ell-1,i}\mathbf{u}$ is defined (only) on $\Omega_{\ell-1,i}^e$; due to (A2) we have that $P_{\ell,i}^{-1} I_{\ell,i}^\Gamma P_{\ell-1,i}\mathbf{u} = P_{\ell,i}^{-1} w_{|\Omega_{\ell,i}^e}$ is well-defined.

LEMMA 3.1. *For the prolongation in* (3.3) *the following holds:*

$$(3.4) \qquad p_\ell P_{\ell-1}^\Gamma \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} = P_\ell^\Gamma \begin{pmatrix} \mathbf{p}_{\ell,1}\mathbf{u}_{H,1} \\ \mathbf{p}_{\ell,2}\mathbf{u}_{H,2} \end{pmatrix} \quad for \quad \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} \in \mathbb{R}^{n_{\ell-1,1}+n_{\ell-1,2}}.$$

*Proof.* Take $\begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} \in \mathbb{R}^{n_{\ell-1,1}+n_{\ell-1,2}}$ and

$$u_H := P_{\ell-1}^\Gamma \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} = \mathcal{R}_{\ell-1} \begin{pmatrix} P_{\ell-1,1}\mathbf{u}_{H,1} \\ P_{\ell-1,2}\mathbf{u}_{H,2} \end{pmatrix},$$

where in the last equality we used (2.7a). Hence, from (3.3) and (2.7b) we get

$$p_\ell P_{\ell-1}^\Gamma \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} = p_\ell u_H = \mathcal{R}_\ell I_\ell^\Gamma \begin{pmatrix} P_{\ell-1,1}\mathbf{u}_{H,1} \\ P_{\ell-1,2}\mathbf{u}_{H,2} \end{pmatrix}$$

$$= P_\ell^\Gamma \begin{pmatrix} P_{\ell,1}^{-1} I_{\ell,1}^\Gamma P_{\ell-1,1}\mathbf{u}_{H,1} \\ P_{\ell,2}^{-1} I_{\ell,2}^\Gamma P_{\ell-1,2}\mathbf{u}_{H,2} \end{pmatrix} = P_\ell^\Gamma \begin{pmatrix} \mathbf{p}_{\ell,1}\mathbf{u}_{H,1} \\ \mathbf{p}_{\ell,2}\mathbf{u}_{H,2} \end{pmatrix},$$

which proves the result. $\qquad\square$

As a direct corollary we obtain that the matrix representation of the prolongation in (3.3), i.e., $\mathbf{p}_\ell = (P_\ell^\Gamma)^{-1} p_\ell P_{\ell-1}^\Gamma$, is given by

$$(3.5) \qquad \mathbf{p}_\ell \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} = \begin{pmatrix} \mathbf{p}_{\ell,1}\mathbf{u}_{H,1} \\ \mathbf{p}_{\ell,2}\mathbf{u}_{H,2} \end{pmatrix} \quad for \quad \begin{pmatrix} \mathbf{u}_{H,1} \\ \mathbf{u}_{H,2} \end{pmatrix} \in \mathbb{R}^{n_{\ell-1,1}+n_{\ell-1,2}}.$$

This is the prolongation used in the multigrid solver. It has the following interesting (and at first sight maybe surprising) property. The prolongation in matrix representation only requires $\mathbf{p}_{\ell,i}\mathbf{u}_{H,i}$, $i = 1, 2$, where $\mathbf{p}_{\ell,i}$ is the prolongation corresponding to $V_{\ell,i}$. This $\mathbf{p}_{\ell,i}$ does *not* depend on $\Gamma_{\ell-1}$ or $\Gamma_\ell$ and is just the simple interpolation used in a standard setting of nested finite element spaces, without any cutting procedures involved. For the matrix-vector representation of the discrete problems we use another basis as the one in (2.6), which will be introduced in the next section. The matrix representation of the prolongation (in the basis (2.6)) can easily be transformed to this other basis.

**3.2. Smoother.** Before we explain the smoother we first introduce the stiffness matrix corresponding to the discretization methods described in section 2. In XFEM one usually chooses a basis of the space $V_\ell^\Gamma$ that differs from the one in (2.6). For the matrix-vector representation of the unfitted finite element discretization we also use this alternative basis, which we now introduce. Recall that $(\varphi_{\ell,i}^k)_{1 \le k \le n_{\ell,i}}$ denotes the standard nodal basis in $V_{\ell,i}$; cf. (2.4). The vertex corresponding to $\varphi_{\ell,i}^k$ is denoted by $x_{k,i}$. Let $\mathcal{V}_i$ be the set of those vertices in $\Omega_{\ell,i}$ that correspond to exactly two degrees of freedom in $V_\ell^\Gamma$ (namely, in $V_{\ell,1}$ and $V_{\ell,2}$). We define

$$(3.6) \qquad B_i := \{\, \varphi_{\ell,i}^k \mid x_{k,i} \in \Omega_{\ell,i} \,\}, \quad B_i^\Gamma := \{\, \varphi_{\ell,i}^k|_{\Omega \setminus \Omega_{\ell,i}} \mid x_{k,i} \in \mathcal{V}_i \,\}, \quad i = 1, 2.$$

Note that $B_1 \cup B_2$ is the nodal basis in the standard finite element space $V_\ell$ (2.6). The disjoint union $B_1 \cup B_2 \cup B_1^\Gamma \cup B_2^\Gamma$ is a basis of $V_\ell^\Gamma$, which we call XFEM basis. The corresponding finite element isomorphism is denoted by $\tilde{P}_\ell^\Gamma : \mathbb{R}^{n_{\ell,1}+n_{\ell,2}} \to V_\ell^\Gamma$. It is easy to transform between this XFEM basis and the one given in (2.6). For the matrix-vector representation of the discretizations we use the XFEM basis. Hence, the stiffness matrix $A_\ell \in \mathbb{R}^{(n_{\ell,1}+n_{\ell,2}) \times (n_{\ell,1}+n_{\ell,2})}$ for (Nitsche) is defined by (with $\langle \cdot, \cdot \rangle$ the Euclidean scalar product)

$$(3.7) \qquad \langle A_\ell \mathbf{u}, \mathbf{v} \rangle = a_h(\tilde{P}_\ell^\Gamma \mathbf{u}, \tilde{P}_\ell^\Gamma \mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n_{\ell,1}+n_{\ell,2}},$$

and similarly for (P-Nitsche) and ($\mu$-Nitsche).

*Remark* 4. Clearly there is an issue related to "small cuts." In [29] for the case of the XFEM basis it is shown that the poor conditioning of this basis with respect to the $L^2$-scalar product can be eliminated by a simple scaling. More precisely, let $M_\ell$ be the mass matrix formed by the $L^2(\Omega)$ scalar product of the XFEM basis functions and $D_\ell := \mathrm{diag}(M_\ell)$. Then $D_\ell^{-1} M_\ell$ has a *uniformly bounded spectral condition number*, where the uniformity holds both with respect to the refinement level $\ell$ and the way in which the triangulation is intersected by the (approximate) interface $\Gamma_\ell$. Such a robustness property w.r.t. the location of $\Gamma_\ell$ has not been proved for the diagonally scaled stiffness matrix. However, in practice one often observes that the condition number of the diagonally scaled stiffness matrix is also robust w.r.t. small cuts; cf. section 4.2. Despite this possibly very poor conditioning (of the unscaled matrix), we conclude from our numerical experiments that the multigrid solver introduced below has a contraction number that is small uniformly in the level number $\ell$, the interface location, and (for the method ($\mu$-Nitsche)) the jump in the diffusion coefficient.

Concerning the choice of the smoother we distinguish two problem classes. If the diffusion coefficient ratio is moderate, i.e., $r_\mu = \mathcal{O}(1)$, we use a *standard Gauss–Seidel smoother*. From the experiments we see that this yields satisfactory results. If, however, this quotient is increased we observe a deterioration of the rate of convergence.

For the case of large coefficient jumps $r_\mu \gg 1$ we propose a modification in which an additional local error damping procedure is included that acts *only close to the interface*. One obvious possibility, used also in multigrid methods in other settings, is to apply additional local smoothing (cf. [8, Chapter 5], [17, 19]), e.g., block Gauss–Seidel with blocks corresponding to local patches close to the interface. Alternatively one could add extra Gauss-Seidel smoothing iterations at grid points close to the interface. From experiments we see that this then results in a fairly efficient multigrid solver, also for larger diffusion quotient ratios. We propose, however, another local correction procedure which results in a solver with even better performance. This smoother, which we call "Gauss–Seidel with interface correction" (GS-IC), is based on a simple domain decomposition approach and is explained below.

**Gauss–Seidel with interface correction.** The smoother (on level $\ell$) needs only information from the discretization on level $\ell$. To simplify notation, in the remainder of this section we delete the level index $\ell$. The set $\mathcal{V}_i = \{x_{1,i}, \ldots, x_{m_i,i}\}$ (cf. (3.6)) contains those vertices in $\Omega_i$ that correspond to exactly two degrees of freedom. Let $\mathcal{I}^\Gamma \subset \{1, \ldots, n_1 + n_2\}$ be the subset which contains the (global) numbering of the unknowns corresponding to the vertices in $\cup_{i=1,2}\mathcal{V}_i$ and $R^\Gamma : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{|\mathcal{I}^\Gamma|}$ the matrix representation of the corresponding injection. Note that $|\mathcal{I}^\Gamma| = 2(m_1 + m_2)$. With that we define the *local interface matrix*

$$(3.8) \qquad A^\Gamma = R^\Gamma A \, (R^\Gamma)^T \in \mathbb{R}^{|\mathcal{I}^\Gamma| \times |\mathcal{I}^\Gamma|},$$

which is the part of the full stiffness matrix $A$ corresponding to the degrees of freedom "close to" the interface. We propose a local interface correction smoother in the spirit of a multiplicative Schwarz smoother, where first a standard smoothing, e.g., Gauss–Seidel, is performed on the entire domain followed by a correction for the interface unknowns.

Systems with the matrix $A^\Gamma$ can be solved (approximately) with arithmetic costs that are (very) low compared to the costs of the Gauss–Seidel iteration in step 2. This will be further addressed in section 3.3.

**3.3. Analysis of local interface matrix $A^\Gamma$.** In the case of moderate or large coefficient jumps we propose to use the GS-IC smoother. For moderate coefficient jumps the discretization methods (Nitsche) and (P-Nitsche) yield satisfactory results. Below, in section 3.3.1 it is shown that for these methods the spectral condition number of the diagonally preconditioned matrix $A^\Gamma$ is bounded *independent of the level $\ell$ and of the location of $\Gamma$ in the triangulation*. From numerical experiments (cf. section 4) we see that a fixed (independent of $\ell$) low number of (preconditioned) conjugate gradient (CG) iterations is sufficient for the *approximate* solution of the linear system in step 4 of GS-IC. Since the number of unknowns in the $A^\Gamma$-system is much lower than the total number of unknowns, $|\mathcal{I}^\Gamma| \ll n_1 + n_2$, and $A^\Gamma$ has the same sparsity pattern as $A$, we thus obtain a very efficient realization of the GS-IC smoother. The condition number of the diagonally preconditioned matrix $A^\Gamma$ depends on the size of the coefficient jump.

For very large $r_\mu$ values we need (too) many CG iterations. In such a case the discretization method ($\mu$-Nitsche) is significantly better than the other two, and for the linear system resulting from this method we propose to use the multigrid solver with GS-IC smoother, and in step 4 we apply a *sparse direct solver*. This still is an efficient smoother as further explained in section 3.3.2.

**3.3.1. Conditioning of matrix $A^\Gamma$.** For an analysis of conditioning properties of the matrix $A^\Gamma \in \mathbb{R}^{|\mathcal{I}^\Gamma| \times |\mathcal{I}^\Gamma|}$ it is convenient to introduce the corresponding representation on the finite element space $V^\Gamma$ (we delete the level index $\ell$). To each $x_{k,i} \in \mathcal{V}_i$ there corresponds one standard nodal finite element function and one cut finite element basis function. These are denoted by $\varphi_i^k$ and $\varphi_{i,\Gamma}^k := (\varphi_i^k)_{|(\Omega \setminus \Omega_i)}$, respectively ($1 \le k \le m_i$). We introduce the (local) subspaces of $V^\Gamma$:

$$\mathcal{W}_0 := \text{span}\{\, \varphi_i^k \mid k = 1, \ldots, m_i, \ \ i = 1, 2 \,\},$$

$$(3.9) \qquad V_i^{\text{ex}} := \text{span}\{\, \varphi_{i,\Gamma}^k \mid k = 1, \ldots, m_i \,\}, \ \ i = 1, 2, \ \ \mathcal{W}_1 := V_1^{\text{ex}} \oplus V_2^{\text{ex}},$$

$$V_{\text{loc}} := \mathcal{W}_0 \oplus \mathcal{W}_1 \subset V^\Gamma.$$

For the finite element isomorphism w.r.t. the XFEM basis $\tilde{P}^\Gamma : \mathbb{R}^{n_1+n_2} \to V^\Gamma$ we have, per construction, that its restriction to the subspace $V_{\text{loc}}$, denoted by

$\hat{P}^\Gamma : \mathbb{R}^{|\mathcal{I}^\Gamma|} \to V_{\mathrm{loc}}$, is an isomorphism. The local matrix $A^\Gamma$ in (3.8) is the matrix representation of $a_h(\cdot, \cdot)$ restricted to $V_{\mathrm{loc}}$:

$$\langle A^\Gamma \mathbf{u}, \mathbf{v} \rangle = a_h(\hat{P}^\Gamma \mathbf{u}, \hat{P}^\Gamma \mathbf{v}) \quad \text{for all} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^{|\mathcal{I}^\Gamma|}.$$

For the analysis of the spectral condition number of $A^\Gamma$ we will use the key property (explained in more detail below) that the splitting $V_{\mathrm{loc}} = \mathcal{W}_0 \oplus \mathcal{W}_1$ is a *stable splitting* w.r.t. $a_h(\cdot, \cdot)$. The restrictions of $\hat{P}^\Gamma$ to the subspace $\mathcal{W}_s$ are denoted by $\hat{P}_s$, $s = 0, 1$:

$$\hat{P}_0 : \mathbb{R}^{m_1 + m_2} \to \mathcal{W}_0, \quad \hat{P}_0 \hat{\mathbf{u}} := \sum_{i=1}^{2} \sum_{k=1}^{m_i} \hat{u}_i^k \varphi_i^k,$$

$$\hat{P}_1 : \mathbb{R}^{m_1 + m_2} \to \mathcal{W}_1, \quad \hat{P}_1 \tilde{\mathbf{u}} := \sum_{i=1}^{2} \sum_{k=1}^{m_i} \tilde{u}_i^k \varphi_{i,\Gamma}^k.$$

The unknowns in $\mathcal{I}^\Gamma$ are numbered accordingly, taking first the unknowns corresponding to $\mathcal{W}_0$ and then those corresponding to $\mathcal{W}_1$. This is denoted by the partitioning $\mathbb{R}^{|\mathcal{I}^\Gamma|} \ni \mathbf{u} = \left( \begin{smallmatrix} \hat{\mathbf{u}} \\ \tilde{\mathbf{u}} \end{smallmatrix} \right)$. We write $V_{\mathrm{loc}} \ni u = \hat{P}^\Gamma \mathbf{u} = \hat{P}_0 \hat{\mathbf{u}} + \hat{P}_1 \tilde{\mathbf{u}} =: \hat{u} + \tilde{u} \in \mathcal{W}_0 + \mathcal{W}_1$ for the splitting of a finite element function. Also the matrix $D := \mathrm{diag}(A^\Gamma)$ is partitioned accordingly, i.e.,

$$D = \begin{pmatrix} D_0 & \emptyset \\ \emptyset & D_1 \end{pmatrix}, \quad \text{and} \quad \langle D\mathbf{u}, \mathbf{u} \rangle = \langle D_0 \hat{\mathbf{u}}, \hat{\mathbf{u}} \rangle + \langle D_1 \tilde{\mathbf{u}}, \tilde{\mathbf{u}} \rangle.$$

We introduce a representation of the Jacobi preconditioner $D_s$ in the subspace $\mathcal{W}_s$, i.e., $B_s : \mathcal{W}_s \to \mathcal{W}_s$ such that

$$(B_s w, z)_{L^2} := \langle D_s \hat{P}_s^{-1} w, \hat{P}_s^{-1} z \rangle \quad \text{for all} \quad w, z \in \mathcal{W}_s, \quad s = 0, 1.$$

For arbitrary $\mathbf{u} \in \mathbb{R}^{|\mathcal{I}^\Gamma|}$, $\mathbf{u} \neq 0$, we then have

$$(3.10) \qquad \frac{\langle A^\Gamma \mathbf{u}, \mathbf{u} \rangle}{\langle D\mathbf{u}, \mathbf{u} \rangle} = \frac{a_h(u, u)}{(B_0 \hat{u}, \hat{u})_{L^2} + (B_1 \tilde{u}, \tilde{u})_{L^2}}.$$

We first consider the standard Nitsche bilinear form $a_h(\cdot, \cdot)$ used in (Nitsche). In [25], for the two-dimensional case $d = 2$ the following *stable splitting property* is derived: there exists a constant $K > 0$, independent of the discretization level $\ell$ and of how the triangulation is intersected by $\Gamma$ such that

$$(3.11) \qquad K\big(a_h(\hat{u}, \hat{u}) + a_h(\tilde{u}, \tilde{u})\big) \leq a_h(\hat{u} + \tilde{u}, \hat{u} + \tilde{u}) \quad \text{for all } \hat{u} \in \mathcal{W}_0, \tilde{u} \in \mathcal{W}_1.$$

It is expected (cf. the numerical experiments for $d = 3$ given in [25]) that this result is also valid for $d = 3$. However, a proof of (3.11) for the case $d = 3$ is not known, yet. Based on (3.11) we obtain

$$(3.12) \qquad \frac{\langle A^\Gamma \mathbf{u}, \mathbf{u} \rangle}{\langle D\mathbf{u}, \mathbf{u} \rangle} \sim \frac{a_h(\hat{u}, \hat{u}) + a_h(\tilde{u}, \tilde{u})}{(B_0 \hat{u}, \hat{u})_{L^2} + (B_1 \tilde{u}, \tilde{u})_{L^2}},$$

where the $\sim$ is used to denote that there are inequalities between the terms on both sides with strictly positive constants independent of the discretization $\ell$ and of how the triangulation is intersected by $\Gamma$. It is easy to show (cf. Lemma 4.2 in [25]) that

the Jacobi preconditioner and $a_h(\cdot, \cdot)$ are spectrally equivalent on $\mathcal{W}_1$ (in fact even on the two subspaces $V_i^{ex}$ of $\mathcal{W}_1$):

$$(3.13) \qquad \frac{a_h(\tilde{u}, \tilde{u})}{B_1(\tilde{u}, \tilde{u})_{L^2}} \sim 1 \quad \text{for all} \ \ \tilde{u} \in \mathcal{W}_1.$$

The proof of this result in [25] holds for both $d = 2$ and $d = 3$. It remains to derive a similar spectral equivalence on $\mathcal{W}_0$. Such a result is given in the following lemma.

LEMMA 3.2. *The spectral equivalence*

$$(3.14) \qquad \frac{a_h(\hat{u}, \hat{u})}{B_0(\hat{u}, \hat{u})_{L^2}} \sim 1 \quad \text{for all} \ \ \hat{u} \in \mathcal{W}_0$$

*holds.*

*Proof.* Take $w \in \mathcal{W}_0$. We write $w = \sum_{i=1}^{2} \sum_{k=1}^{m_i} w_i^k \varphi_i^k =: \sum_{i,k} w_i^k \varphi_i^k$ and $\hat{P}_0^{-1} w = \mathbf{w} = (w_i^k)_{1 \le i \le 2, 1 \le k \le m_i}$. Because $w$ is continuous we have $a_h(w, w) = a(w, w) = \int_\Omega \mu (\nabla w)^2 \, dx$. Due to the finite overlap property of finite element nodal functions we have

$$a_h(w, w) \lesssim \sum_{i,k} (w_i^k)^2 \int_\Omega \mu \big(\nabla \varphi_i^k\big)^2 dx = \langle D_0 \mathbf{w}, \mathbf{w} \rangle,$$

and thus

$$(3.15) \qquad a_h(w, w) \lesssim (B_0 w, w)_{L^2}$$

holds, which yields the "$\lesssim$" inequality in (3.14). The constant in $\lesssim$ depends only on shape regularity properties of the triangulation. For every vertex $x_{k,i} \in \mathcal{V}_i$, $i = 1, 2$, $k = 1, \ldots, m_i$, we can select a simplex $T_{k,i} \in \mathcal{T}$ such that 1. $x_{k,i}$ is a vertex of $T_{k,i}$, 2. at least one vertex of $T_{k,i}$ is *not* in $\mathcal{V}_1 \cup \mathcal{V}_2$, and 3. $T_{k,i} \cap T_{m,j} = \emptyset$ if $(k, i) \ne (m, j)$. See Figure 3.2 for illustration. Below we use (different) strictly positive constants, denoted by $c$, which depend only on the shape regularity of the triangulation. Note that for $w \in \mathcal{W}_0$ the local linear function $w_{|T_{k,i}}$ has a zero value at at least one vertex of $T_{k,i}$, and thus

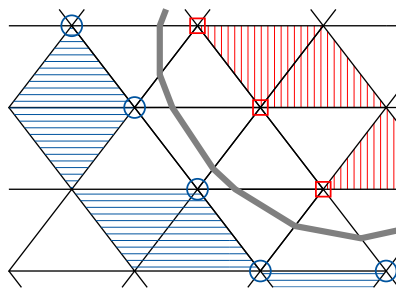$$\nabla w_{|T_{k,i}} \cdot \nabla w_{|T_{k,i}} \ge c \Big( \frac{w(x_{k,i})}{h} \Big)^2.$$



FIG. 3.2. *Node sets $\mathcal{V}_1$ ◯ and $\mathcal{V}_2$ ☐ with corresponding simplices $T_{k,1}$ ⊟ and $T_{k,2}$ ⊞.*

Hence,

$$a_h(w, w) = \int_\Omega \mu \nabla w \cdot \nabla w \, dx \geq \mu_{\min} \sum_{i,k} \int_{T_{k,i}} \nabla w \cdot \nabla w \, dx$$

(3.16)
$$\geq \mu_{\min} c h^{-2} \sum_{i,k} |T_{i,k}| w(x_{k,i})^2$$

$$\geq \mu_{\min} c h^{d-2} \sum_{i,k} w(x_{k,i})^2 = \mu_{\min} c h^{d-2} \langle \mathbf{w}, \mathbf{w} \rangle.$$

For the diagonal matrix $D_0$ we have

$$(D_0)_{(k,i),(k,i)} = a_h(\varphi_i^k, \varphi_i^k) \leq \mu_{\max} \int_\Omega (\nabla \varphi_i^k)^2 \, dx \leq \mu_{\max} c h^{d-2},$$

and thus

$$(B_0 w, w)_{L^2} = \langle D_0 \mathbf{w}, \mathbf{w} \rangle \leq \mu_{\max} c h^{d-2} \langle \mathbf{w}, \mathbf{w} \rangle.$$

Combining this with (3.16) we obtain the "$\gtrsim$" estimate in (3.14). $\qquad\square$

We summarize the results derived above in the following corollary.

COROLLARY 3.3. *For $d = 2$ the local interface matrix $A^\Gamma$ resulting from the standard Nitsche method* (Nitsche) *has the property*

(3.17)
$$\frac{\langle A^\Gamma \mathbf{u}, \mathbf{u} \rangle}{\langle D \mathbf{u}, \mathbf{u} \rangle} \sim 1 \quad \textit{for all} \ \ \mathbf{u} \in \mathbb{R}^{|\mathcal{I}^\Gamma|}, \ \mathbf{u} \neq 0.$$

*The constants in the spectral equivalence $\sim$ are independent of the discretization level $\ell$ and of how the triangulation is intersected by $\Gamma$. Assume (3.11) holds for $d = 3$. Then (3.17) holds also for $d = 3$.*

*Proof.* From (3.10) and (3.11) we obtain the equivalence (3.12). The result (3.17) follows from this by using (3.13) and (3.14). $\qquad\square$

We now address the parameter-free Nitsche method (P-Nitsche) with bilinear form $\tilde{a}_h(\cdot, \cdot)$. In [23] it is shown that $\tilde{a}_h(\cdot, \cdot) \sim a_h(\cdot, \cdot)$ on the unfitted finite element space $V_\ell^\Gamma$ holds. Hence, the conditioning result in Corollary 3.3 also holds for $\tilde{a}_h(\cdot, \cdot)$ (parameter-free Nitsche).

These results on the conditioning of $A^\Gamma$ are illustrated in section 4.3.2; cf. Table 4.10 and Table 4.11.

*Remark* 5. In the analysis above, all arguments are valid both for $d = 2$ and $d = 3$, except the stable splitting property in (3.11), which is derived in [25] only for $d = 2$. This stable splitting property, however, is a key ingredient in the proof of (3.17). The result in (3.11) is equivalent to the strengthened Cauchy inequality

(3.18)
$$a_h(\hat{u}, \tilde{u}) \leq (1 - K) a_h(\hat{u}, \hat{u})^{\frac{1}{2}} a_h(\tilde{u}, \tilde{u})^{\frac{1}{2}} \quad \text{for all } \hat{u} \in \mathcal{W}_0, \ \tilde{u} \in \mathcal{W}_1.$$

Denote the local domain consisting of all elements intersected by the interface by $\Omega_\ell^\Gamma := \cup_{T \in \mathcal{T}_\ell^\Gamma} T$. Note that $\mathrm{supp}(\tilde{u}) \subset \Omega_\ell^\Gamma$ for all $\tilde{u} \in \mathcal{W}_1$. Using this property it follows that the (local) strengthened Cauchy inequality (3.18) is equivalent to the (global) one where the local space $\mathcal{W}_0$ is replaced by the global space $V_\ell \supset \mathcal{W}_0$. The strengthened Cauchy inequality with $\mathcal{W}_1$ and this global space $V_\ell$ is derived (for $d = 2$) in Theorem 3.1 in [25].

The constants in (3.17) are independent of the discretization level $\ell$ and of the interface position, but these constants may depend on the diffusion coefficient ratio $r_\mu$.

The dependence of (some of) these constants on $r_\mu$ is intricate. Numerical experiments (section 4.3.2) reveal that for $r_\mu \lesssim 10^2$ the diagonally preconditioned CG method is an efficient solver, with a (essentially) level independent rate of convergence. For larger $r_\mu$ values we recommend to use a sparse direct solver; cf. the next section.

**3.3.2. Direct sparse solver for $A^\Gamma$.** We discuss the computational complexity of solving $A^\Gamma \mathbf{x} = R^\Gamma \mathbf{r}$ in step 4 via a sparse direct solver. We consider $d = 3$ and assume that $\Gamma$ is simply connected (as in the examples in section 4). The number of unknowns in the system $A_\ell \mathbf{u} = \mathbf{b}_\ell$ is of order $\mathcal{O}(h_\ell^{-3})$ and the computational costs of one Gauss–Seidel smoothing iteration is also $\mathcal{O}(h_\ell^{-3})$. The local interface matrix $A^\Gamma$ has dimension $|\mathcal{I}^\Gamma| \times |\mathcal{I}^\Gamma|$ with $|\mathcal{I}^\Gamma| = \mathcal{O}(h_\ell^{-2})$. Assume that the vertices $x_{k,i}$, $i = 1,2$, $k = 1,\ldots,m_i$, are ordered with a breadth first search traversal of the corresponding triangulation (starting from an arbitrary vertex). The standard finite element mass matrix (with this ordering), i.e., $M \in \mathbb{R}^{\frac{1}{2}|\mathcal{I}^\Gamma| \times \frac{1}{2}|\mathcal{I}^\Gamma|}$ with entries $M_{(k,i),(m,j)} = \int_\Omega \varphi_i^k \varphi_j^m \, dx$, has a bandwidth of size $\mathcal{O}(h_\ell^{-1})$. In the system $A^\Gamma \mathbf{x} = R^\Gamma \mathbf{r}$ we use a blocking of the unknowns with at each vertex $x_{k,i}$ the two unknowns corresponding to $\varphi_i^k$ and $\varphi_{i,\Gamma}^k$; cf. (3.9). Hence, the matrix $A^\Gamma$ has a block structure with

$$(A^\Gamma)_{(k,i),(m,j)} = \begin{pmatrix} a_h(\varphi_j^m, \varphi_i^k) & a_h(\varphi_{j,\Gamma}^m, \varphi_i^k) \\ a_h(\varphi_j^m, \varphi_{i,\Gamma}^k) & a_h(\varphi_{j,\Gamma}^m, \varphi_{i,\Gamma}^k) \end{pmatrix}.$$

Since for all three discretization methods $M_{(k,i),(m,j)} = 0$ (disjoint supports of the basis functions) implies $(A^\Gamma)_{(k,i),(m,j)} = 0$, it follows that $A^\Gamma$ has a bandwidth of size $\mathcal{O}(h_\ell^{-1})$. Thus the computation of a sparse Cholesky factorization of $A^\Gamma$ requires $\mathcal{O}(h_\ell^{-3})$ arithmetic operations, which is of the same order of magnitude as the costs for one Gauss–Seidel iteration. A numerical illustration of this optimal computational complexity for the sparse direct solver is given in section 4.3.2; cf. Table 4.12.

**4. Numerical experiments.** In the numerical experiments we use the domain $\Omega = [0,2]^3$ and two different types of interfaces $\Gamma$ (cf. Figure 4.1):
- planar interface: $\Gamma_{\text{plan}}(x_\Gamma) := \{(x,y,z)^T \in \Omega \mid x = x_\Gamma\}$,
- spherical interface. For given $\mathbf{m} \in \mathbb{R}^3$, $r > 0$, we define the level set function $\phi(\mathbf{x}) := \|\mathbf{x} - \mathbf{m}\|_2^2 - r^2$ and $\Gamma_{\text{spher}}(r, \mathbf{m}) := \{\mathbf{x} \in \Omega \mid \phi(\mathbf{x}) = 0\}$. Unless stated otherwise, we use $\mathbf{m} = (1.03, 1.02, 1.01)$, $r = 0.413$.

The initial mesh $\mathcal{T}_0$ is quasi-uniform with $6 \cdot 4^3$ tetrahedra, and $\mathcal{T}_\ell$ for $\ell > 1$ is obtained by uniform refinement of $\mathcal{T}_{\ell-1}$, i.e., $|\mathcal{T}_\ell| = 8 \cdot |\mathcal{T}_{\ell-1}|$. For the spherical interface the interface approximation in the experiments is given by $\Gamma_\ell = \{\mathbf{x} \in \Omega \mid (I_{\ell+1}\phi)(\mathbf{x}) = 0\}$ with $I_{\ell+1}$ the nodal interpolation in the standard finite element space $V_{\ell+1}$. This iso$P_2$ choice (instead of the linear interpolation on the current mesh $\mathcal{T}_\ell$) is
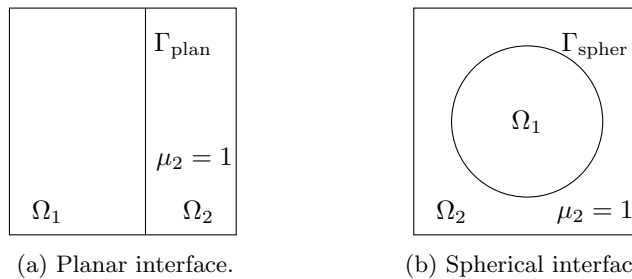


(a) Planar interface.  (b) Spherical interface.

FIG. 4.1. *Sketch of the problem domains.*

motivated by the fact that the level set function $\phi$ is a $P_2$ finite element function. We obtain very similar results if we use $I_\ell\phi$ instead of $I_{\ell+1}\phi$. This interface approximation satisfies (A1), (A2) in Assumption 1.

In all experiments the stiffness matrix corresponds to the XFEM basis (3.6), and the prolongation matrix (3.5) is transformed to this basis.

**4.1. Discretization error.** For testing the proposed discretizations (Nitsche), (P-Nitsche), and ($\mu$-Nitsche) we consider the case with $\Gamma_{\mathrm{spher}}$ with a prescribed solution

$$u^* = \alpha(\mathbf{x})\left(\|\mathbf{x} - \mathbf{m}\|_2^2 - r^2\right) \ \ \text{with} \ \ \alpha(\mathbf{x}) := \begin{cases} \mu_2 & \text{for } \mathbf{x} \in \Omega_1, \\ \mu_1 & \text{for } \mathbf{x} \in \Omega_2. \end{cases}$$

We use $\mu_2 = 1$ for all experiments. Unless stated otherwise, we choose $\mu_1 < 1$.

Hence, $u^*$ fulfills (1.1a)–(1.1d) when setting $u_D = u^*$ and $f = -6\mu_1\mu_2$. Tables 4.1–4.3 show, for various choices of $\mu_1$, the convergence results on different refinement levels for the three discretizations. As measure for the convergence we use the $L^2$-error $e(\ell) := \|u_{h_\ell} - u^*\|_{0,\Omega}$.

TABLE 4.1
$L^2$-errors for $u_{h_\ell}$ obtained by (Nitsche) with $\lambda_N = 10$ for various choices of $\mu_1$ with corresponding experimentally determined orders of convergence (e.o.c.). Spherical interface.

| $\ell$ | $r_\mu = 1.\bar{1}$ $e(\ell)$ | e.o.c. | $r_\mu = 10$ $e(\ell)$ | e.o.c. | $r_\mu = 10^3$ $e(\ell)$ | e.o.c. | $r_\mu = 10^5$ $e(\ell)$ | e.o.c. |
|---|---|---|---|---|---|---|---|---|
| 0 | 3.27E-01 | | 4.83E-02 | | 4.38E-02 | | 4.41E-02 | |
| 1 | 8.20E-02 | 2.00 | 1.69E-02 | 1.51 | 4.07E-02 | 0.11 | 4.42E-02 | 0.00 |
| 2 | 2.05E-02 | 2.00 | 3.98E-03 | 2.09 | 2.51E-02 | 0.69 | 3.46E-02 | 0.35 |
| 3 | 5.14E-03 | 2.00 | 9.47E-04 | 2.07 | 8.58E-03 | 1.55 | 2.01E-02 | 0.78 |
| 4 | 1.28E-03 | 2.00 | 2.27E-04 | 2.06 | 1.72E-03 | 2.31 | 1.05E-02 | 0.94 |

TABLE 4.2
$L^2$-errors for $u_{h_\ell}$ obtained by (P-Nitsche) for various choices of $\mu_1$ with corresponding convergence order. Spherical interface.

| $\ell$ | $r_\mu = 1.\bar{1}$ $e(\ell)$ | e.o.c. | $r_\mu = 10$ $e(\ell)$ | e.o.c. | $r_\mu = 10^3$ $e(\ell)$ | e.o.c. | $r_\mu = 10^5$ $e(\ell)$ | e.o.c. |
|---|---|---|---|---|---|---|---|---|
| 0 | 3.26E-01 | | 4.40E-02 | | 4.21E-02 | | 4.41E-02 | |
| 1 | 8.19E-02 | 2.00 | 1.36E-02 | 1.69 | 3.27E-02 | 0.36 | 4.37E-02 | 0.01 |
| 2 | 2.05E-02 | 2.00 | 3.45E-03 | 1.98 | 1.07E-02 | 1.60 | 3.35E-02 | 0.38 |
| 3 | 5.14E-03 | 2.00 | 8.66E-04 | 2.00 | 2.16E-03 | 2.31 | 1.80E-02 | 0.89 |
| 4 | 1.28E-03 | 2.00 | 2.16E-04 | 2.00 | 3.79E-04 | 2.51 | 7.29E-03 | 1.31 |

TABLE 4.3
$L^2$-errors for $u_{h_\ell}$ obtained by ($\mu$-Nitsche) with stabilization parameter $\varepsilon_g = 0.1$ and $\lambda_N = 10$ for various choices of $\mu_1$ with corresponding convergence order. Spherical interface.

| $\ell$ | $r_\mu = 1.\bar{1}$ $e(\ell)$ | e.o.c. | $r_\mu = 10$ $e(\ell)$ | e.o.c. | $r_\mu = 10^3$ $e(\ell)$ | e.o.c. | $r_\mu = 10^5$ $e(\ell)$ | e.o.c. |
|---|---|---|---|---|---|---|---|---|
| 0 | 3.26E-01 | | 4.59E-02 | | 2.49E-02 | | 2.49E-02 | |
| 1 | 8.44E-02 | 1.95 | 1.65E-02 | 1.47 | 1.32E-02 | 0.91 | 1.32E-02 | 0.91 |
| 2 | 2.08E-02 | 2.02 | 4.00E-03 | 2.05 | 3.14E-03 | 2.08 | 3.14E-03 | 2.08 |
| 3 | 5.18E-03 | 2.01 | 9.47E-04 | 2.08 | 7.24E-04 | 2.12 | 7.23E-04 | 2.12 |
| 4 | 1.29E-03 | 2.01 | 2.28E-04 | 2.05 | 1.70E-04 | 2.09 | 1.70E-04 | 2.09 |

The methods (Nitsche) and (P-Nitsche) show optimal order of convergence only for modest contrast in the viscosities, whereas ($\mu$-Nitsche) is robust for the whole range (that we considered). We obtain similar results when varying $\mu_1$ instead of $\mu_2$ and keeping $\mu_2$ fixed, and therefore those tables are not included here.

**4.2. Condition number dependence.** We consider the stiffness matrix arising from (Nitsche) with the XFEM basis (3.7) and fix $r_\mu = 2$. We first vary the position of the spherical interface $\Gamma_{\mathrm{spher}}$ by changing the origin $\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{e}$ with constant vectors $\mathbf{m}_0 = (1.03, 1.02, 1.01)^T$, $\mathbf{e} = (1, 1, 1)^T$, and variable $\delta \in [0, 0.04]$. By moving the sphere through the mesh, different situations of "small" cuts occur. Figure 4.2 shows, for different values of $\delta$, the spectral condition number (denoted by $\kappa_2(\cdot)$) of the stiffness matrix $A_4$ and of the diagonally scaled stiffness matrix $D_4^{-1}A_4$ on the finest level $\ell = 4$.

We observe that $\kappa_2(A_4)$ obtains very large values and shows an irregular behavior depending on the cut situations, whereas $\kappa_2(D_4^{-1}A_4)$ remains constant which is in accordance with [25]. Within the smoother of the multigrid method a diagonal scaling is used, which might explain the robustness (with respect to the interface location) of the multigrid solver applied to $A_\ell$, despite the very large values and irregular behavior of $\kappa_2(A_\ell)$; cf. Table 4.9. Similar results are obtained for (P-Nitsche). For ($\mu$-Nitsche) the condition number $\kappa_2(A_4)$ is stable independent of the relative interface position thanks to the ghost penalty stabilization.

Now we fix $\Gamma = \Gamma_{\mathrm{spher}}$ for $\delta = 0$ and compare the condition numbers for varying $\mu_1$ and refinement level $\ell$. As the refinements are uniform, in Table 4.4 we can observe that $\kappa_2(D_\ell^{-1}A_\ell)$ scales with $h_\ell^{-2}$ which is typical for Poisson problems. This condition number is (for the considered range of $h_\ell$ values) not proportional to $r_\mu$ but seems to stagnate for large $r_\mu$. Similar results are obtained for (P-Nitsche). Note that for the
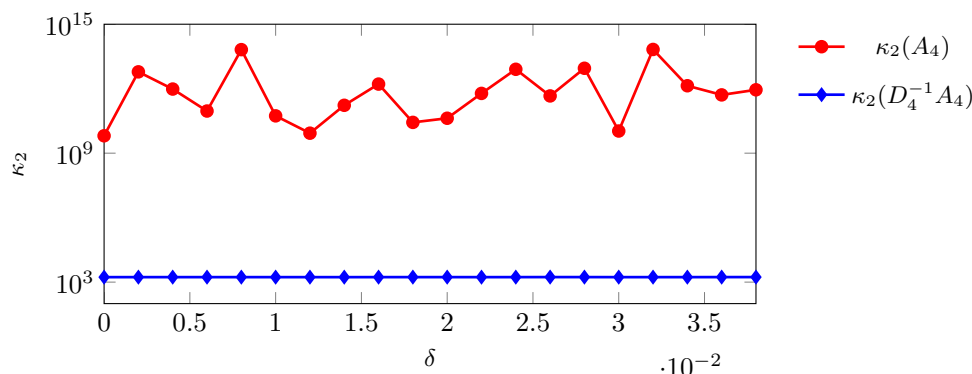


FIG. 4.2. *Condition number for system matrix (unscaled and diagonally scaled) on refinement level $\ell = 4$ with fixed $r_\mu = 2$ for $\Gamma_{spher}$ with varying origin.*

TABLE 4.4
*Condition numbers for $\kappa_2(D_\ell^{-1}A_\ell)$ of (Nitsche) for variation in $\mu$ and $\ell$.*

| $\ell$ | $r_\mu = 2$ | 10 | $10^3$ | $10^5$ | $10^7$ |
|---|---|---|---|---|---|
| 1 | $2.75 \cdot 10^1$ | $6.05 \cdot 10^1$ | $2.98 \cdot 10^2$ | $3.11 \cdot 10^2$ | $3.11 \cdot 10^2$ |
| 2 | $1.04 \cdot 10^2$ | $1.02 \cdot 10^2$ | $7.55 \cdot 10^2$ | $9.56 \cdot 10^2$ | $9.58 \cdot 10^2$ |
| 3 | $4.23 \cdot 10^2$ | $4.10 \cdot 10^2$ | $1.64 \cdot 10^3$ | $3.56 \cdot 10^3$ | $3.60 \cdot 10^3$ |
| 4 | $1.70 \cdot 10^3$ | $1.64 \cdot 10^3$ | $2.44 \cdot 10^3$ | $1.37 \cdot 10^4$ | $1.45 \cdot 10^4$ |

values $r_\mu = 10^5$, $r_\mu = 10^7$, the results for $\kappa_2(D_\ell^{-1} A_\ell)$ are less relevant, because both (Nitsche) and (P-Nitsche) are not $r_\mu$ robust and these methods should not be used for discretization in case of very large $r_\mu$ values. For the $r_\mu$ robust method ($\mu$-Nitsche) there is a significantly stronger dependence of $\kappa_2(D_\ell^{-1} A_\ell)$ on $r_\mu$. This may be due to the fact that for large $r_\mu$ values this discretization is "closer" to the continuous operator, for which the conditioning clearly deteriorates for larger $r_\mu$ values.

**4.3. Multigrid solver.** In this section we present results of experiments with the multigrid method explained in section 3. Every multigrid iteration consists of one V-cycle with 2 pre- and postsmoothing steps. The iteration is stopped once a relative residual (in the Euclidean norm) of $10^{-8}$ is reached. Depending on the ratio $r_\mu$ of the diffusion coefficient, different discretization methods and smoothing strategies are applied. We refer to the Introduction (section 1) for a motivation concerning the different discretization methods and to section 3.2 for a discussion on the different smoothers; cf. also Table 1.1 for an overview.

**4.3.1. Multigrid solver for small coefficient jumps.** We restrict to the case of small/moderate ratios $r_\mu \lesssim 10^2$ with (Nitsche) and (P-Nitsche) as discretization methods.

**Exact interface.** We first consider the problem (1.1) with a planar interface. In this case there is no error in the interface approximation and $\Gamma_\ell = \Gamma_{\text{plan}}$. Therefore the finite element spaces are nested, i.e., $V_{\ell-1}^{\Gamma_{\ell-1}} \subset V_\ell^{\Gamma_\ell}$ for $\ell = 1, 2, \ldots$, and the prolongation corresponds to simple injection based on the nesting of the spaces.

We conduct the experiments with $x_\Gamma = 1.321$ for $\Gamma_{\text{plan}}$, the right-hand-side $f = xyz$, and boundary data $u_D = 0$ in problem (1.1). Table 4.5 shows the multigrid iteration numbers for (Nitsche) with $\lambda_N = 10$ and varying $\mu$-ratios.

We observe that for moderate coefficient jumps the multigrid iteration numbers are essentially independent of the discretization level. However, the (Nitsche) discretization requires the penalty parameter $\lambda_N$ which has a strong influence on the iteration numbers. The influence of $\lambda_N$ is illustrated in Table 4.6 for a fixed discretization level $\ell = 2$ and the moderate coefficient ratio $r_\mu = 2$.

Due to the strong dependence of the multigrid iteration numbers on $\lambda_N$, we repeat the experiment from Table 4.5 but change the discretization from (Nitsche) to (P-Nitsche) where the penalty term is implicitly determined.

The iteration numbers in Table 4.7 are almost identical to the ones in Table 4.5 for the coefficient jumps $r_\mu \in \{1.\bar{1}, 2\}$ and are more robust for larger contrast. Due to the

TABLE 4.5

(Nitsche)*: Multigrid iteration numbers for various coefficient ratios and fixed $x_\Gamma = 1.321$.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | $r_\mu = 2$ | $r_\mu = 10$ | $r_\mu = 100$ |
|---|---|---|---|---|
| 1 | 8 | 10 | 31 | 72 |
| 2 | 10 | 10 | 24 | 127 |
| 3 | 11 | 11 | 36 | 297 |
| 4 | 11 | 11 | 25 | 219 |

TABLE 4.6

*Multigrid iteration numbers for different $\lambda_N$ and fixed $x_\Gamma = 1.321$, $\ell = 2$, and $r_\mu = 2$.*

| $\lambda_N$ | 1 | 10 | 20 | 100 | 1000 |
|---|---|---|---|---|---|
| iterations | 11 | 10 | 13 | 43 | 143 |

TABLE 4.7

(P-Nitsche): *Multigrid iteration numbers for various coefficient ratios and fixed $x_\Gamma = 1.321$.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | $r_\mu = 2$ | $r_\mu = 10$ | $r_\mu = 100$ |
|---|---|---|---|---|
| 1 | 8 | 8 | 18 | 93 |
| 2 | 10 | 10 | 12 | 77 |
| 3 | 10 | 10 | 16 | 144 |
| 4 | 11 | 11 | 15 | 99 |

benefits of the parameter-free discretization method (P-Nitsche), in the experiments below we restrict to (P-Nitsche) and ($\mu$-Nitsche).

**Approximate interface.** In the above experiments the interface $\Gamma$ was planar, and hence on the different levels $\ell$ we have $\Gamma_\ell = \Gamma$. Now we consider the test problem described above with $\Gamma_{\mathrm{spher}}$. In this case the interface approximations on different levels vary, i.e., $\Gamma_{\ell-1} \neq \Gamma_\ell$, and therefore the finite element spaces are not nested: $V_{\ell-1}^{\Gamma_{\ell-1}} \not\subset V_\ell^{\Gamma_\ell}$. Table 4.8 shows the multigrid iteration numbers for varying coefficient ratios.

We observe that the results are similar to the exact interface case in Table 4.7. This is due to the use of a canonical prolongation operator $p_\ell : V_{\ell-1}^{\Gamma_{\ell-1}} \to V_\ell^{\Gamma_\ell}$.

To further investigate the multigrid robustness with respect to the interface position, we fix the coefficient ratio $r_\mu = 2$. We vary the interface position for $\Gamma_{\mathrm{spher}}$ by changing the origin of the sphere $\mathbf{m} = \mathbf{m}_0 + \delta \mathbf{e}$ with $\mathbf{m}_0$ and $\mathbf{e}$ as in section 4.2 and variable $\delta$. The results for $\ell = 1, \ldots, 4$ are displayed in Table 4.9.

The multigrid iteration numbers are (almost) the same for every interface position, and we conclude that the rate of convergence of the method is not sensitive with respect to the interface location.

**4.3.2. Multigrid solver for large coefficient jumps.** As seen in the previous section, for increasing coefficient ratios $r_\mu$ the multigrid iteration numbers deteriorate. Therefore, we repeat the experiment from Table 4.8, i.e., for the (P-Nitsche) discretization of the test problem with the spherical interface, but instead of a standard Gauss–Seidel smoother we use the GS-IC smoother from Algorithm 1. Table 4.10

TABLE 4.8

(P-Nitsche): *Multigrid iteration numbers for various coefficient ratios and $\Gamma_{spher}$.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | $r_\mu = 2$ | $r_\mu = 10$ | $r_\mu = 100$ |
|---|---|---|---|---|
| 1 | 8 | 8 | 18 | 89 |
| 2 | 10 | 10 | 17 | 101 |
| 3 | 11 | 11 | 14 | 88 |
| 4 | 11 | 11 | 16 | 118 |

TABLE 4.9

(P-Nitsche): *Multigrid iteration numbers for various interface positions. Spherical interface.*

| $\ell$ | $\delta = 0$ | $\delta = 0.1$ | $\delta = 0.2$ | $\delta = 0.3$ |
|---|---|---|---|---|
| 1 | 8 | 9 | 8 | 8 |
| 2 | 10 | 10 | 10 | 10 |
| 3 | 11 | 11 | 11 | 11 |
| 4 | 11 | 11 | 11 | 11 |

**Algorithm 1.** Gauss–Seidel with interface correction (GS-IC).

---

1: **function** $S_\Gamma^C(\mathbf{x}^n, \mathbf{b}, A)$

2:     $\mathbf{x}^{n+\frac{1}{2}} = S(\mathbf{x}^n, \mathbf{b}, A)$                    ▷ perform standard smoothing

3:     $\mathbf{r} = A\mathbf{x}^{n+\frac{1}{2}} - \mathbf{b}$                    ▷ new residual

4:     $\mathbf{x}^{n+1} = \mathbf{x}^{n+\frac{1}{2}} - (R^\Gamma)^T (A^\Gamma)^{-1} R^\Gamma \mathbf{r}$                    ▷ interface update

5:     **return** $\mathbf{x}^{n+1}$

6: **end function**

---

TABLE 4.10
(P-Nitsche): *Multigrid iteration numbers for various coefficient ratios with GS-IC smoother. The maximum numbers of CG iterations required to solve the system (with relative tolerance $10^{-2}$) with $A^\Gamma$ are displayed in brackets. Spherical interface.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | 2 | 10 | $10^2$ | $10^3$ | $10^4$ | $10^5$ |
|---|---|---|---|---|---|---|---|
| 1 | 7 (10) | 7 (10) | 7 (15) | 7 (31) | 7 | 7 | 7 |
| 2 | 9 (10) | 9 (10) | 9 (15) | 9 (37) | 9 | 9 | 9 |
| 3 | 10 (11) | 10 (11) | 10 (15) | 10 (38) | 11 | 17 | 18 |
| 4 | 10 (12) | 10 (12) | 10 (15) | 10 (42) | 12 | 21 | 29 |

shows the multigrid iteration numbers. In brackets for small to medium contrast we put the maximum number of diagonally preconditioned CG iterations to reach a relative tolerance of $10^{-2}$ for the solution of $A^\Gamma \mathbf{x} = R^\Gamma \mathbf{r}$. For a larger contrast it is more efficient to use a direct solver for the interface matrix, as too many CG iterations would be required (e.g., 496 iterations for $\ell = 4$ and $r_\mu = 10^5$).

We observe that multigrid iteration numbers are robust for much larger coefficient ratios than with a standard Gauss–Seidel smoother. For small to medium contrast, the iterations required to solve the systems with the interface matrix $A^\Gamma$ are very low and almost constant w.r.t. the refinement level $\ell$. This confirms the uniform (in $\ell$) condition number bounds derived in section 3.3.1. For larger $\mu$-contrast the multigrid method is less robust. Note, however, that for large $r_\mu$ values ($r_\mu \gtrsim 10^3$) the discretization (P-Nitsche) is not satisfactory, due to suboptimal convergence behavior.

*Remark* 6. Solving the systems with $A^\Gamma$ approximately with a tolerance of $10^{-2}$ for the relative residual yields (almost) the same multigrid iteration numbers as a direct solve.

In Table 4.11 we show condition numbers of the diagonally scaled interface matrix needed within GS-IC smoother. These confirm the theoretical findings given in

TABLE 4.11
*Condition number of the diagonally scaled interface matrix $A^\Gamma$ in Algorithm 1 for various levels $\ell$ and different $r_\mu$.*

| $\ell$ | (P-Nitsche): $\kappa_2(D^{-1}A_\ell^\Gamma)$ | | (Nitsche): $\kappa_2(D^{-1}A_\ell^\Gamma)$ | |
|---|---|---|---|---|
| | $r_\mu = 10^2$ | $r_\mu = 10^4$ | $r_\mu = 10^2$ | $r_\mu = 10^4$ |
| 1 | 300 | 1343 | 213 | 309 |
| 2 | 224 | 3533 | 235 | 916 |
| 3 | 243 | 6420 | 251 | 2904 |
| 4 | 268 | 8225 | 342 | 6911 |

TABLE 4.12

*Number of nonzero entries in $A_\ell$ ($\mathcal{N}(A_\ell)$) and comparison with number of nonzero entries of sparse Cholesky factor in $A_\ell^\Gamma = L_\ell L_\ell^T$.*

| $\ell$ | $|\mathcal{N}(A_\ell)|$ | $|\mathcal{N}(L_\ell)|/|\mathcal{N}(A_\ell)|$ |
|---|---|---|
| 1 | 6,835 | 0.71 |
| 2 | 56,038 | 0.58 |
| 3 | 466,267 | 0.45 |
| 4 | 3,824,281 | 0.33 |

section 3.3. For $r_\mu = 10^2$ the almost constant small condition number is reflected by almost constant small iteration numbers of the interface CG solver; on the other hand, for $r_\mu = 10^4$, the condition numbers are much larger, and for solving the $A^\Gamma$ linear systems (with tolerance $10^{-2}$) a very high number of CG iterations is needed. For such high contrast cases we recommend to use a sparse direct solver.

In Table 4.12 we give numbers for the fill-in in the lower triangular matrix $L_\ell$ of the sparse Cholesky decomposition $A_\ell^\Gamma = L_\ell L_\ell^T$ of the local interface matrix. These results confirm that the direct solver has optimal computational efficiency; cf. section 3.3.2.

**Coefficient stable and interface stabilized discretization.** As seen in the experiments above, the proposed multigrid method is only robust up to a certain contrast in the diffusion coefficient. The deterioration in the multigrid convergence rate may be related to the lack of robustness of the discretization method (P-Nitsche). Clearly, for large coefficient jumps it is better to use a discretization method that is robust with respect to the coefficient ratio, such as ($\mu$-Nitsche). We apply the multigrid method, with GS-IC smoother, to the linear system resulting from the discretization method ($\mu$-Nitsche). Results are shown in Table 4.13. We observe a very robust behavior of the multigrid method, with iteration numbers that are essentially independent of the value of $r_\mu$.

So far we only investigated the parameter configuration $\mu_2 = 1$ and $\mu_1 < 1$, i.e., "soft inclusion." We repeat the previous experiment with the reversed setting $\mu_1 = 1$ and $\mu_2 < 1$, i.e., "hard inclusion;" see Table 4.14.

TABLE 4.13

*($\mu$-Nitsche): Multigrid iteration numbers with GS-IC smoother for various coefficient ratios ($\mu_2 = 1$, soft inclusion), with stabilization parameter $\varepsilon_g = 0.1$ and $\lambda_N = 10$. Spherical interface.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | $r_\mu = 10$ | $r_\mu = 10^3$ | $r_\mu = 10^5$ | $r_\mu = 10^7$ |
|---|---|---|---|---|---|
| 1 | 7 | 7 | 7 | 7 | 7 |
| 2 | 9 | 9 | 9 | 9 | 9 |
| 3 | 10 | 10 | 10 | 10 | 10 |
| 4 | 11 | 11 | 11 | 11 | 11 |

TABLE 4.14

*($\mu$-Nitsche): Multigrid iteration numbers with GS-IC smoother for various coefficient ratios ($\mu_1 = 1$, hard inclusion), with stabilization parameter $\varepsilon_g = 0.1$ and $\lambda_N = 10$. Spherical interface.*

| $\ell$ | $r_\mu = 1.\bar{1}$ | $r_\mu = 10$ | $r_\mu = 10^3$ | $r_\mu = 10^5$ | $r_\mu = 10^7$ |
|---|---|---|---|---|---|
| 1 | 7 | 7 | 7 | 7 | 7 |
| 2 | 9 | 9 | 10 | 11 | > 100 |
| 3 | 10 | 11 | 12 | 13 | > 100 |
| 4 | 11 | 11 | 12 | 14 | > 100 |

TABLE 4.15

($\mu$-Nitsche): *Multigrid iteration numbers with GS-IC smoother for $\Gamma_{spher}$ with stabilization parameter $\varepsilon_g = 0.1$ on fixed $\ell = 2$ and for different coefficient ratios and various $\lambda_N$.*

| $\lambda_N$ | 1 | 10 | 20 | 100 | 1000 |
|---|---|---|---|---|---|
| Iterations ($r_\mu = 10$) | 12 | 9 | 9 | 9 | 9 |
| Iterations ($r_\mu = 10^5$) | div | 9 | 9 | 9 | 9 |

Similar to Table 4.13 we observe that method is essentially independent of $r_\mu$, except for the case $r_\mu = 10^7$. There we observe that the relative residuals stagnate around $10^{-7}$.

*Remark* 7. For the reversed experiment in Table 4.14 we observed that the multigrid iteration numbers are also robust for $r_\mu = 10^7$, when in the stopping criterion the Euclidian norm is replaced by the norm $\| \cdot \|_D$, where $D$ is the diagonal of the system matrix. We did not investigate this effect any further.

Finally we repeat the experiment from Table 4.6 for the ($\mu$-Nitsche) discretization with the GS-IC smoother (however for the spherical interface $\Gamma_{\text{spher}}$ instead of the planar one). Results are shown in Table 4.15.

The method diverges for $r_\mu = 10^5$ with stabilization parameter value $\lambda_N = 1$, which is probably due to the fact that the discretization is unstable for $\lambda_N$ too small. For all other cases, we observe that the multigrid solver is robust with respect to variations in $\lambda_N$. The robustness is due to the use of the (robust) GS-IC smoother. Applying the latter to the (Nitsche) discretization yields similar results.

*Remark* 8. We obtain almost identical multigrid iteration numbers for all three discretizations when using the Galerkin approach for the coarse grid matrices $A_{\ell-1} = \mathbf{r}_\ell A_\ell \mathbf{p}_\ell$. It can be shown that these Galerkin coarse grid matrices have no extra fill-in compared to the case of direct discretization.

**5. Conclusions.** We considered the discrete problems resulting from three known unfitted finite element discretizations of the Poisson interface problem ((Nitsche), (P-Nitsche), and ($\mu$-Nitsche)). The main topic of the paper is the development of an efficient multigrid method for this class of discrete problems. A tailor-made prolongation operator for the unfitted finite element space was constructed based on the injection of standard finite element spaces on locally extended subdomains. Using level-dependent interface approximations for the discretization of the coarse grid matrices, which can be seen as a kind of quadrature error, did not affect the multigrid convergence. For the case of large contrast in the diffusion coefficients a novel interface smoother has been presented. For small contrast a standard Gauss–Seidel smoothing is already sufficient to obtain an efficient multigrid method. For the Nitsche method (Nitsche) a strong dependence of the multigrid convergence rate on the penalty parameter $\lambda_N$ was observed, and therefore the parameter-free discretization (P-Nitsche) is preferred. Even though the condition number of the resulting stiffness matrix strongly depends on the interface position (for the case without stabilization), the resulting multigrid method did not show any dependence on that. For the case of large coefficient jumps the stabilized discretization ($\mu$-Nitsche) turns out to be better (as known from the literature). The application of the smoother with local interface correction results in a robust multigrid method. Properties of the interface smoother, concerning conditioning and fill-in in the Cholesky decomposition, are derived. These results imply optimal arithmetic costs for one multigrid iteration.

## REFERENCES

[1] L. ADAMS AND T. P. CHARTIER, *New geometric immersed interface multigrid solvers*, SIAM J. Sci. Comput., 25 (2004), pp. 1516–1533.

[2] L. ADAMS AND Z. LI, *The immersed interface/multigrid methods for interface problems*, SIAM J. Sci. Comput., 24 (2002), pp. 463–479.

[3] C. ANNAVARAPU, M. HAUTEFEUILLE, AND J. DOLBOW, *A robust Nitsche's formulation for interface problems*, Comput. Methods Appl. Mech. Engrg., 225 (2012), pp. 44–54.

[4] B. AYUSO DE DIOS, M. HOLST, Y. ZHU, AND L. ZIKATANOV, *Multilevel preconditioners for discontinuous Galerkin approximations of elliptic problems with jump coefficients*, Math. Comp., 83 (2014), pp. 1083–1120.

[5] S. BADIA AND F. VERDUGO, *Robust and scalable domain decomposition solvers for unfitted finite element methods*, J. Comput. Appl. Math., 344 (2018), pp. 740–759.

[6] N. BARRAU, R. BECKER, E. DUBACH, AND R. LUCE, *A robust variant of NXFEM for the interface problem*, C. R. Math. Acad. Sci. Paris, 350 (2012), pp. 789–792.

[7] L. BERGER-VERGIAT, H. WAISMAN, B. HIRIYUR, R. TUMINARO, AND D. KEYES, *Inexact Schwarz-algebraic multigrid preconditioners for crack problems modeled by extended finite element methods*, Internat. J. Numer. Methods Engrg., 90 (2012), pp. 311–328.

[8] J. H. BRAMBLE, *Multigrid Methods*, CRC Press, Boca Raton, FL, 1993.

[9] E. BURMAN, *La pénalisation fantôme*, C. R. Math. Acad. Sci. Paris, 348 (2010), pp. 1217–1220.

[10] E. BURMAN, S. CLAUS, P. HANSBO, M. LARSON, AND A. MASSING, *CutFEM: Discretizing geometry and partial differential equations*, Internat. J. Numer. Methods Engrg., 104 (2015), pp. 472–501.

[11] E. BURMAN AND P. ZUNINO, *Numerical approximation of large contrast problems with the unfitted Nitsche method*, in Frontiers in Numerical Analysis - Durham 2010, J. Blowey and M. Jensen, eds., Springer-Verlag, Berlin, 2011, pp. 227–282.

[12] T. CHEN AND J. STRAIN, *Piecewise-polynomial discretization and Krylov-accelerated multigrid for elliptic interface problems*, J. Comput. Phys., 227 (2008), pp. 7503–7542.

[13] F. DE PRENTER, C. VERHOOSEL, AND H. VAN BRUMMELEN, *Preconditioning immersed isogeometric finite element methods with application to flow problems*, Comput. Methods Appl. Mech. Engrg., 348 (2019), pp. 604–631.

[14] F. DE PRENTER, C. V. VERHOOSEL, G. VAN ZWIETEN, AND E. H. VAN BRUMMELEN, *Condition number analysis and preconditioning of the finite cell method*, Comput. Methods Appl. Mech. Engrg., 316 (2017), pp. 297–327.

[15] T.-P. FRIES AND T. BELYTSCHKO, *The extended/generalized finite element method: An overview of the method and its applications*, Internat. J. Numer. Methods Engrg., 84 (2010), pp. 253–304.

[16] A. GERSTENBERGER AND R. TUMINARO, *An algebraic multigrid approach to solve extended finite element method based fracture problems*, Internat. J. Numer. Methods Engrg., 94 (2013), pp. 248–272.

[17] M. GRIEBEL AND P. OSWALD, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70 (1995), pp. 163–180.

[18] A. HANSBO AND P. HANSBO, *An unfitted finite element method, based on Nitsche's method, for elliptic interface problems*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 5537–5552.

[19] B. JANSSEN AND G. KANSCHAT, *Adaptive multilevel methods with local smoothing for $H^1$- and $H^{curl}$-conforming high order finite element methods*, SIAM J. Sci. Comput., 33 (2011), pp. 2095–2114.

[20] G. JO AND D. Y. KWAK, *Geometric multigrid algorithms for elliptic interface problems using structured grids*, Numer. Algorithms, (2018), pp. 1–25.

[21] C. LEHRENFELD, *On a Space-Time Extended Finite Element Method for the Solution of a Class of Two-Phase Mass Transport Problems*, Ph.D. thesis, Universitätsbibliothek der RWTH Aachen, Aachen, Germany, 2015.

[22] C. LEHRENFELD, *High order unfitted finite element methods on level set domains using isoparametric mappings*, Comput. Methods Appl. Mech. Engrg., 300 (2016), pp. 716–733.

[23] C. LEHRENFELD, *Removing the Stabilization Parameter in Fitted and Unfitted Symmetric Nitsche Formulations*, arXiv preprint arXiv:1603.00617, 2016.

[24] C. LEHRENFELD AND A. REUSKEN, *L2-error analysis of an isoparametric unfitted finite element method for elliptic interface problems*, J. Numer. Math., (2016).

[25] C. Lehrenfeld and A. Reusken, *Optimal preconditioners for Nitsche-XFEM discretizations of interface problems*, Numer. Math., 135 (2017), pp. 313–332.

[26] R. J. Leveque and Z. Li, *The immersed interface method for elliptic equations with discontinuous coefficients and singular sources*, SIAM J. Numer. Anal., 31 (1994), pp. 1019–1044.

[27] Z. Li, *A fast iterative algorithm for elliptic interface problems*, SIAM J. Numer. Anal., 35 (1998), pp. 230–254.

[28] J. Nitsche, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Semin. Univ. Hamburg, 36, (1971), pp. 9–15.

[29] A. Reusken, *Analysis of an extended pressure finite element space for two-phase incompressible flows*, Comput. Visual. Sci., 11 (2008), pp. 293–305.

[30] W. L. Wan, *Interface preserving coarsening multigrid for elliptic problems with highly discontinuous coefficients*, Numer. Linear Algebra Appl., 7 (2000), pp. 727–742.