

Arnold Reusken

# Analysis of an extended pressure finite element space for two-phase incompressible flows

the date of receipt and acceptance should be inserted later

*Dedicated to Wolfgang Hackbusch  
on the occasion of his 60th birthday*

**Abstract** We consider a standard model for incompressible two-phase flows in which a localized force at the interface describes the effect of surface tension. If a level set (or VOF) method is applied then the interface, which is implicitly given by the zero level of the level set function, is in general not aligned with the triangulation that is used in the discretization of the flow problem. This non-alignment causes severe difficulties w.r.t. the discretization of the localized surface tension force and the discretization of the flow variables. In cases with large surface tension forces the pressure has a large jump across the interface. In standard finite element spaces, due to the non-alignment, the functions are continuous across the interface and thus not appropriate for the approximation of the discontinuous pressure. In many simulations these effects cause large oscillations of the velocity close to the interface, so-called spurious velocities. In [6] it is shown that an extended finite element space (XFEM) is much better suited for the discretization of the pressure variable. In this paper we derive important properties of the XFEM space. We present (optimal) approximation error bounds and prove that the diagonally scaled mass matrix has a uniformly bounded spectral condition number. Results of numerical experiments are presented that illustrate properties of the XFEM space.

**Mathematics Subject Classification (2000)** 65M60 · 65N15 · 65N30 · 76D45 · 76T99

**Keywords** extended finite element space · two-phase flow · surface tension · spurious velocities

---

A. Reusken  
Institut für Geometrie und Praktische Mathematik,  
RWTH Aachen, D-52056 Aachen, Germany. E-mail:  
reusken@igpm.rwth-aachen.de

---

## 1 Introduction

In two-phase incompressible flow problems with surface tension the pressure typically is discontinuous across the interface. In interface capturing methods (for example, level set techniques) the approximation of the unknown interface is not aligned with the grid used for the discretization of the flow problem. Due to this standard polynomial finite element spaces ( $P_0$ -discontinuous or  $P_1$ -continuous) have very poor approximation quality when used for the discretization of the pressure variable. In a previous paper [6] it is shown that an extended finite element space (XFEM) is much better suited for the approximation of the pressure. In this paper we study this XFEM space and introduce a modified XFEM space that has better LBB-stability properties when combined with a standard  $P_2$ -finite element space for the velocity approximation. In this introduction we briefly explain a few major issues related to the finite element discretization of incompressible two-phase flow problems. To simplify the presentation we restrict ourselves to *stationary* problems. We emphasize, however, that the (modified) XFEM space is appropriate both for stationary and instationary problems.

We consider a stationary two-phase incompressible flow problem. Let  $\Omega \subset \mathbb{R}^3$  be a convex polyhedral domain containing two different immiscible incompressible phases. The subdomains containing the two phases are denoted by  $\Omega_1$  and  $\Omega_2$  with  $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$  and  $\Omega_1 \cap \Omega_2 = \emptyset$ . We assume that  $\Omega_1$  and  $\Omega_2$  are connected. The interface is denoted by  $\Gamma = \bar{\Omega}_1 \cap \bar{\Omega}_2$  and is assumed to be sufficiently smooth. The standard model for describing incompressible two-phase flows consists of the Navier-Stokes equations in the subdomains with the coupling condition

$$[\boldsymbol{\sigma}\mathbf{n}]_{\Gamma} = \tau\mathcal{K}\mathbf{n}$$

at the interface, i. e., the surface tension balances the jump of the normal stress at the interface. We use the

notation  $[\cdot]_\Gamma$  for the jump across  $\Gamma$ ,  $\mathbf{n} = \mathbf{n}_\Gamma$  is the unit normal at the interface  $\Gamma$  (pointing from  $\Omega_1$  into  $\Omega_2$ ),  $\tau$  the constant surface tension coefficient,  $\mathcal{K}$  the curvature of  $\Gamma$  and  $\boldsymbol{\sigma}$  the stress tensor defined by

$$\boldsymbol{\sigma} = -p\mathbf{I} + \mu\mathbf{D}(\mathbf{u}), \quad \mathbf{D}(\mathbf{u}) = \nabla\mathbf{u} + (\nabla\mathbf{u})^T,$$

with  $p = p(\mathbf{x})$  the pressure,  $\mathbf{u} = \mathbf{u}(\mathbf{x})$  the velocity and  $\mu$  the viscosity. We assume continuity of  $\mathbf{u}$  across the interface and use homogeneous Dirichlet boundary conditions, i. e.,  $\mathbf{u} = 0$  on  $\partial\Omega$ . For a weak formulation of this problem (as in, for example, [3, 10, 11, 13, 14]) we introduce the spaces

$$\mathbf{V} := H_0^1(\Omega)^3,$$

$$Q := L_0^2(\Omega) = \{q \in L^2(\Omega) \mid \int_\Omega q \, dx = 0\}.$$

The standard  $L^2(\Omega)$  scalar product is denoted by  $(\cdot, \cdot)_0$  and for the Sobolev norm in  $\mathbf{V}$  we use the notation  $\|\cdot\|_1$ . The weak formulation is as follows: determine  $(\mathbf{u}, p) \in \mathbf{V} \times Q$  such that for all  $\mathbf{v} \in \mathbf{V}$  and all  $q \in Q$

$$\begin{aligned} \int_\Omega \mu \mathbf{D}(\mathbf{u}) : \mathbf{D}(\mathbf{v}) \, dx + (\rho \mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v})_0 + (\operatorname{div} \mathbf{v}, p)_0 \\ = (\rho \mathbf{g}, \mathbf{v})_0 + f_\Gamma(\mathbf{v}), \end{aligned} \quad (1)$$

$$(\operatorname{div} \mathbf{u}, q)_0 = 0$$

holds, with

$$f_\Gamma(\mathbf{v}) := \tau \int_\Gamma \mathcal{K} \mathbf{n}_\Gamma \cdot \mathbf{v} \, ds \quad (2)$$

the localized surface tension force and  $\mathbf{D}(\mathbf{u}) : \mathbf{D}(\mathbf{v}) = \operatorname{tr}((\mathbf{D}(\mathbf{u})\mathbf{D}(\mathbf{v})))$ . The functions  $\mu$  and  $\rho$  are strictly positive and piecewise constant in  $\Omega_i$ ,  $i = 1, 2$ , with values  $\mu = \mu_i$ ,  $\rho = \rho_i$  in  $\Omega_i$ . For  $\Gamma$  sufficiently smooth we have  $\sup_{x \in \Gamma} |\mathcal{K}(x)| \leq c < \infty$  and

$$|f_\Gamma(\mathbf{v})| \leq c\tau \int_\Gamma |\mathbf{n}_\Gamma \cdot \mathbf{v}| \, ds \leq c\|\mathbf{v}\|_{L^2(\Gamma)} \leq c\|\mathbf{v}\|_1 \quad (3)$$

for all  $\mathbf{v} \in \mathbf{V}$ .

Thus  $f_\Gamma \in \mathbf{V}'$  holds and hence (1) is well-defined. Under the usual assumptions (cf. [4]) the weak formulation of the Navier-Stokes equations as in (1) has a unique solution. Due to the Laplace-Young law, typically the pressure has a jump across the interface, when surface tension forces are present ( $\tau \neq 0$ ), cf. Remark 1 below. In numerical simulations, this discontinuity and inadequate approximation of the localized surface force term often lead to strong unphysical oscillations of the velocity vector at the interface, so called spurious velocities. In [6] it is shown that these spurious solutions can be avoided to a large extent if a modified Laplace-Beltrami discretization, as explained in [6] (and analyzed in [7]), is used *and* if for the Galerkin discretization of the pressure an *extended finite element space* (XFEM) is used.

Such extended finite element spaces are introduced by Belytschko, cf. [1, 9] in the context of elasticity problems. This XFEM space is also used in [9, 8] for interface problems.

In this paper we derive important properties of this XFEM space related to its approximation quality and stability of the basis used in this space. Furthermore, we introduce a modified XFEM space that is obtained from the XFEM space by deleting finite element basis functions which have a “very small” support. We will quantify what we mean by “very small”. This modified space has the same approximation quality as the original XFEM space but better stability properties. As far as we know the approximation and stability properties of the (modified) XFEM space that are derived in this paper are not known in the literature. We also briefly address the issue of LBB stability of the  $P_2$ -XFEM pair, i.e. velocity is discretized by piecewise quadratics and pressure is taken from the (modified) XFEM space.

To present the main ideas and the motivation for the use of an XFEM space we further simplify (1) and consider a Stokes problem with a constant viscosity ( $\mu_1 = \mu_2 = \mu$  in  $\Omega$ ). We emphasize, however, that the methods that we present are *not* restricted to this simplified problem but apply to (instantaneous) Navier-Stokes equations (1) as well. We introduce the following Stokes problem: find  $(\mathbf{u}, p) \in \mathbf{V} \times Q$  such that

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) &= (\rho \mathbf{g}, \mathbf{v})_0 + f_\Gamma(\mathbf{v}) \quad \text{for all } \mathbf{v} \in \mathbf{V}, \\ b(\mathbf{u}, q) &= 0 \quad \text{for all } q \in Q, \end{aligned} \quad (4)$$

where

$$a(\mathbf{u}, \mathbf{v}) := \int_\Omega \mu \nabla \mathbf{u} \nabla \mathbf{v} \, dx, \quad b(\mathbf{v}, q) = \int_\Omega q \operatorname{div} \mathbf{v} \, dx,$$

with a viscosity  $\mu > 0$  that is constant in  $\Omega$  and  $\rho = \rho_i$  piecewise constant in  $\Omega_i$ ,  $i = 1, 2$ . The unique solution of this problem is denoted by  $(\mathbf{u}^*, p^*) \in \mathbf{V} \times Q$ .

**Remark 1** The problem (4) has a *smooth* velocity solution  $\mathbf{u}^* \in (H^2(\Omega))^3 \cap \mathbf{V}$  and a *piecewise smooth* pressure solution  $p$  with  $p|_{\Omega_i} \in H^1(\Omega_i)$ ,  $i = 1, 2$ , which has a jump across  $\Gamma$ . These smoothness properties can be derived as follows. The curvature  $\mathcal{K}$  is a smooth function (on  $\Gamma$ ). Thus there exist  $\hat{p}_1 \in H^1(\Omega_1)$  such that  $(\hat{p}_1)|_\Gamma = \mathcal{K}$  (in the sense of traces). Define  $\hat{p} \in L^2(\Omega)$  by  $\hat{p} = \hat{p}_1$  in  $\Omega_1$ ,  $\hat{p} = 0$  on  $\Omega_2$ . Note that for all  $\mathbf{v} \in \mathbf{V}$ ,

$$\begin{aligned} f_\Gamma(\mathbf{v}) &= \tau \int_\Gamma \mathcal{K} \mathbf{n}_\Gamma \cdot \mathbf{v} \, ds = \tau \int_\Gamma \hat{p}_1 \mathbf{n}_\Gamma \cdot \mathbf{v} \, ds \\ &= \tau \int_{\Omega_1} \hat{p}_1 \operatorname{div} \mathbf{v} \, dx + \tau \int_{\Omega_1} \nabla \hat{p}_1 \cdot \mathbf{v} \, dx \\ &= \tau \int_\Omega \hat{p} \operatorname{div} \mathbf{v} \, dx + \tau \int_\Omega \hat{\mathbf{g}} \cdot \mathbf{v} \, dx, \end{aligned}$$

with  $\tilde{\mathbf{g}} \in L^2(\Omega)^3$  given by  $\tilde{\mathbf{g}} = \nabla \hat{p}_1$  in  $\Omega_1$ ,  $\tilde{\mathbf{g}} = 0$  on  $\Omega_2$ . Thus  $(\mathbf{u}^*, p^* - \tau \hat{p})$  satisfies the standard Stokes equations

$$\begin{aligned} a(\mathbf{u}^*, \mathbf{v}) + b(\mathbf{v}, p^* - \tau \hat{p}) &= (\rho \mathbf{g} + \tau \tilde{\mathbf{g}}, \mathbf{v})_0 \quad \text{for all } \mathbf{v} \in \mathbf{V}, \\ b(\mathbf{u}^*, q) &= 0 \quad \text{for all } q \in Q. \end{aligned}$$

From regularity results on Stokes equations and the fact that  $\Omega$  is convex we conclude that  $\mathbf{u}^* \in H^2(\Omega) \cap H_0^1(\Omega)$  and  $p^* - \tau \hat{p} \in H^1(\Omega)$ . Thus  $[p^* - \tau \hat{p}]_\Gamma = 0$  (a.e. on  $\Gamma$ ) holds, which implies

$$[p^*]_\Gamma = \tau [\hat{p}]_\Gamma = -\tau \mathcal{K},$$

i.e.,  $p^*$  has a jump across  $\Gamma$  of the size  $\tau \mathcal{K}$ .

**Example 1** A simple example, that is used in the numerical experiments in section 5 is the following. Let  $\Omega := (-1, 1)^3$  and  $\Omega_1$  a sphere with centre at the origin and radius  $r < 1$ . We take  $\mathbf{g} = 0$ . In this case the curvature is constant,  $\mathcal{K} = \frac{2}{r}$ , and the solution of the Stokes problem (4) is given by  $\mathbf{u}^* = 0$  in  $\Omega$ ,  $p^* = \tau \frac{2}{r} + c_0$  on  $\Omega_1$ ,  $p^* = c_0$  on  $\Omega_2$  with a constant  $c_0$  such that  $\int_\Omega p^* dx = 0$ .

We apply a Galerkin conforming finite element discretization to the problem (4). Let  $\{\mathcal{T}_h\}_{h>0}$  be a stable family of consistent (i.e., no hanging nodes) nested triangulations, consisting of tetrahedra. In our experiments these triangulations are locally refined close to the interface  $\Gamma$ , cf. section 5. Let  $\mathbf{V}_h \subset \mathbf{V}$ ,  $Q_h \subset Q$  be a stable pair of finite element spaces. We assume that a piecewise planar surface  $\Gamma_h$  is known, such that

$$\text{dist}(\Gamma, \Gamma_h) \leq c h_\Gamma^2, \quad (5)$$

with  $h_\Gamma$  the size (diameter) of the tetrahedra in the locally refined region that contains the interface. This assumption is reasonable if one uses piecewise quadratic finite elements for the discretization of the level set function, cf. [7]. Note that in general the faces of  $\Gamma_h$  are not aligned with the faces in the tetrahedral triangulation  $\mathcal{T}_h$ . The induced polyhedral approximations of the subdomains  $\Omega_1$  and  $\Omega_2$  are denoted by  $\Omega_{1,h}$  and  $\Omega_{2,h}$ , respectively. Furthermore, we define the piecewise constant approximation of the density  $\rho_h$  by  $\rho_h = \rho_i$  on  $\Omega_{i,h}$ . We assume that for  $\mathbf{v}_h \in \mathbf{V}_h$  the integrals in

$$(\rho_h \mathbf{g}, \mathbf{v}_h)_0 = \rho_1 \int_{\Omega_{1,h}} \mathbf{g} \cdot \mathbf{v}_h dx + \rho_2 \int_{\Omega_{2,h}} \mathbf{g} \cdot \mathbf{v}_h dx$$

are computed exactly. Let  $f_{\Gamma_h}(\mathbf{v}_h)$  be a numerical approximation of  $f_\Gamma(\mathbf{v}_h)$ . The standard Galerkin discretization of (4) is as follows: determine  $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$  such that

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) &= (\rho_h \mathbf{g}, \mathbf{v}_h) + f_{\Gamma_h}(\mathbf{v}_h) \quad \text{for all } \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}_h, q_h) &= 0 \quad \text{for all } q_h \in Q_h. \end{aligned} \quad (6)$$

Using standard finite element error analysis (Strang lemma) we get a discretization error bound, cf. [6]:

**Theorem 1** Let  $(\mathbf{u}^*, p^*)$ ,  $(\mathbf{u}_h, p_h)$  be the solution of (4) and (6), respectively. Then the error bound

$$\begin{aligned} &\mu \|\mathbf{u}_h - \mathbf{u}^*\|_1 + \|p_h - p^*\|_{L^2} \\ &\leq c \left( \mu \inf_{\mathbf{v}_h \in \mathbf{V}_h} \|\mathbf{v}_h - \mathbf{u}^*\|_1 + \inf_{q_h \in Q_h} \|q_h - p^*\|_{L^2} \right. \\ &\quad \left. + \sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{|(\rho \mathbf{g}, \mathbf{v}_h) - (\rho_h \mathbf{g}, \mathbf{v}_h)|}{\|\mathbf{v}_h\|_1} \right. \\ &\quad \left. + \sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{|f_\Gamma(\mathbf{v}_h) - f_{\Gamma_h}(\mathbf{v}_h)|}{\|\mathbf{v}_h\|_1} \right) \end{aligned} \quad (7)$$

holds with a constant  $c$  independent of  $h$ ,  $\mu$  and  $\rho$ .

We comment on the terms occurring in the bound in (7). As explained above (Remark 1), the solution  $\mathbf{u}^*$  of (4) is smooth and thus with standard finite element spaces  $\mathbf{V}_h$  for the velocity (e.g.,  $P_1$  or  $P_2$ ) we obtain  $\inf_{\mathbf{v}_h \in \mathbf{V}_h} \|\mathbf{v}_h - \mathbf{u}^*\|_1 \leq ch$ . Due to (5) we get  $|\text{vol}(\Omega_i) - \text{vol}(\Omega_{i,h})| \leq ch_\Gamma^2$ ,  $i = 1, 2$ , and using this we obtain

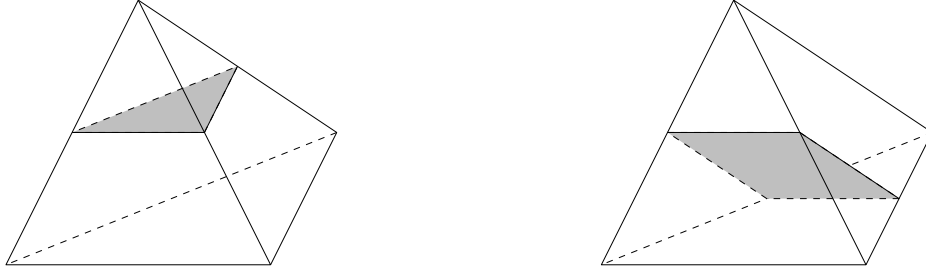
$$\begin{aligned} &|(\rho \mathbf{g}, \mathbf{v}_h)_0 - (\rho_h \mathbf{g}, \mathbf{v}_h)_0| \\ &\leq \sum_{i=1}^2 \rho_i \left| \int_{\Omega_i} \mathbf{g} \cdot \mathbf{v}_h dx - \int_{\Omega_{i,h}} \mathbf{g} \cdot \mathbf{v}_h dx \right| \\ &\leq c(\rho_1 + \rho_2) h_\Gamma \|\mathbf{v}_h\|_1, \end{aligned}$$

and thus an  $\mathcal{O}(h_\Gamma)$  bound for the third term in (7). The remaining two terms in (7) are less easy to handle. In [7] it is shown that a (not so obvious) approximation method based on a Laplace-Beltrami representation results in a functional  $f_{\Gamma_h}(\cdot)$  with an  $\mathcal{O}(h_\Gamma)$  error bound for the last term in (7). This functional  $f_{\Gamma_h}(\cdot)$  is described in section 5. The second term in (7) is discussed in section 2. It is easy to show (cf. [6]) that standard finite element spaces  $Q_h$  (e.g.,  $P_0$  or  $P_1$ ) lead to an error  $\inf_{q_h \in Q_h} \|q_h - p^*\|_{L^2} \sim \sqrt{h_\Gamma}$ . This motivates the use of another pressure finite element space, namely an extended finite element space as explained in section 2, which has optimal approximation properties for functions that are piecewise smooth but discontinuous across  $\Gamma_h$ .

## 2 Extended finite element space

In this section we describe a modified finite element method (for discretization of the pressure variable) that is based on an extension of the standard linear finite element space. This XFEM space has optimal approximation properties for piecewise smooth functions, as is shown in section 3.

In practice the interface  $\Gamma$  is often approximated by an interface capturing method like, for example, a level set method. The level set function is discretized and (an approximation) of the zero level of this discrete level set function is used as an approximation  $\Gamma_h$  of  $\Gamma$ . Clearly, in general this  $\Gamma_h$  changes if the mesh is refined. To avoid many technicalities caused by this variation of  $\Gamma_h$  we



**Fig. 1** Planar intersections of  $\Gamma$  and  $T$ .

assume that  $\Gamma_h = \Gamma$  for all  $h$ . Furthermore, we assume that for all  $T \in \mathcal{T}_h$  the intersection  $\Gamma_h \cap T$  is either empty or a planar segment that does not contain any vertices of  $T$ . In the latter case only two situations can occur, namely  $\Gamma_h \cap T$  is either a triangle or a quadrilateral, cf. Fig. 2.

A two-dimensional case in which such simplifying assumptions are fulfilled is shown in Fig. 2. For  $m \geq 1$  let  $H^m(\Omega_1 \cup \Omega_2)$  denote the Sobolev space of functions, for which  $u|_{\Omega_i} \in H^m(\Omega_i)$ ,  $i = 1, 2$ , holds. We use the notation  $\|u\|_{m, \Omega_1 \cup \Omega_2}^2 = \|u\|_{m, \Omega_1}^2 + \|u\|_{m, \Omega_2}^2$  for  $u \in H^m(\Omega_1 \cup \Omega_2)$ . The  $L^2(\Omega)$  scalar product and corresponding norm are denoted by  $(\cdot, \cdot)_0$  and  $\|\cdot\|_0$ , respectively.

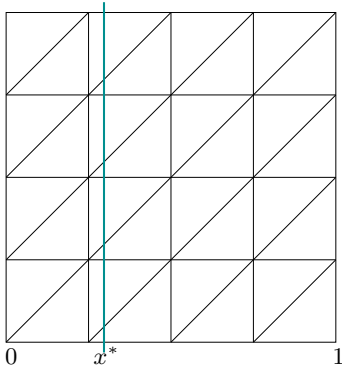
We introduce the standard linear finite element space

$$V = V_h = \{v \in C(\Omega) \mid v|_T \in \mathcal{P}_1 \text{ for all } T \in \mathcal{T}_h\}.$$

For the approximation of functions in  $H^m(\Omega_1 \cup \Omega_2)$ ,  $m \geq 1$ , that are discontinuous across  $\Gamma$  (in trace sense) the finite element space  $V$  is not suitable. In general, for  $u \in H^m(\Omega_1 \cup \Omega_2)$ , one can not expect a better bound than

$$\inf_{v \in V} \|u - v\|_0 \leq c\sqrt{h} \|u\|_{m, \Omega_1 \cup \Omega_2},$$

cf. [6]. To improve this poor approximation quality we extend the space  $V$  by adding functions that can represent



Uniform refinement:  $h_k = 2^{-k}$

$\Gamma$  located at  $x = x^*$  with  $x^*$  irrational.

**Fig. 2** 2D example of a family of triangulations and interface that satisfy the assumptions.

discontinuities across  $\Gamma$ . For this we first introduce some further notation. To simplify this notation we do not express the dependence on  $h$  in our notation (for example,  $V$  instead of  $V_h$ ). Let  $\mathcal{V} = \{x_k\}_{k \in \mathcal{I}}$ ,  $\mathcal{I} = \{1, \dots, n\}$ , be the set of all vertices in the triangulation  $\mathcal{T}_h$ . Note that due to the assumptions on  $\Gamma$  we have  $x_k \notin \Gamma$  for all  $k$ . The nodal basis in  $V$  is denoted by  $\{\phi_k\}_{k \in \mathcal{I}}$ . Let  $\Omega_\Gamma$  be the collection of all tetrahedra that are intersected by  $\Gamma$ , i.e.,  $\Omega_\Gamma = \cup\{T \in \mathcal{T}_h \mid T \cap \Gamma \neq \emptyset\}$ . Let  $R_i : L^2(\Omega) \rightarrow L^2(\Omega)$ ,  $i = 1, 2$ , be *restriction operators*:

$$R_i v = \begin{cases} v|_{\Omega_i} & \text{on } \Omega_i \\ 0 & \text{on } \Omega \setminus \Omega_i \end{cases}$$

(in  $L^2$ -sense). We define

$$\Omega_i^e := \cup\{T \mid T \cap \Omega_i \neq \emptyset\}, \quad i = 1, 2,$$

$$V_i^e := \{v \in C(\Omega_i^e) \mid v|_T \in \mathcal{P}_1 \text{ for all } T \in \Omega_i^e\}, \quad i = 1, 2.$$

We need *extension operators*  $E_i : R_i V \rightarrow V_i^e$ ,  $i = 1, 2$ , given by

$$E_i R_i v = v|_{\Omega_i^e} \quad \text{for } v \in V, \quad i = 1, 2.$$

The standard *nodal interpolation operator* is denoted by  $I : C(\overline{\Omega}_1) \cup C(\overline{\Omega}_2) \rightarrow V$ ,  $(Iv)(x_k) = v(x_k)$  for all  $x_k \in \mathcal{V}$ . We introduce subsets of  $\mathcal{I}$  for which the corresponding basis functions have a nonzero intersection with  $\Gamma$ :

$$\mathcal{I}_1^\Gamma := \{k \in \mathcal{I} \mid x_k \in \Omega_2 \text{ and } \text{supp}(\phi_k) \cap \Gamma \neq \emptyset\}$$

$$\mathcal{I}_2^\Gamma := \{k \in \mathcal{I} \mid x_k \in \Omega_1 \text{ and } \text{supp}(\phi_k) \cap \Gamma \neq \emptyset\}.$$

Corresponding spaces are defined by

$$V_i^\Gamma := \text{span}\{R_i \phi_k \mid k \in \mathcal{I}_i^\Gamma\}, \quad i = 1, 2.$$

The *extended finite element space* is given by

$$V^\Gamma := R_1 V \oplus R_2 V. \quad (8)$$

The following properties will be useful in our analysis:

$$R_i R_j v = 0 \quad \text{for all } v \text{ and } i \neq j, \quad (9)$$

$$R_i R_i v = R_i v \quad \text{for all } v, \quad i = 1, 2, \quad (10)$$

$$R_1 v + R_2 v = v \quad \text{for all } v \in V^\Gamma, \quad (11)$$

$$R_i V^\Gamma = R_i V, \quad i = 1, 2, \quad (12)$$

$$E_i R_i v = v|_{\Omega_i^e} \quad \text{if } v|_{\Omega_i^e} \in V_i^e, \quad (13)$$

$$v(x_k) = 0 \quad \text{for all } v \in V_i^\Gamma, \quad i = 1, 2, \text{ and all } x_k \in \mathcal{V}. \quad (14)$$

We derive another characterization of  $V^\Gamma$ .

**Theorem 2** *The following holds:*

$$V^\Gamma = V \oplus V_1^\Gamma \oplus V_2^\Gamma.$$

Furthermore, in the decomposition

$$v = w + \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k^{(1)} R_1 \phi_k + \sum_{k \in \mathcal{I}_2^\Gamma} \beta_k^{(2)} R_2 \phi_k, \quad (15)$$

with  $v \in V^\Gamma$ ,  $w \in V$ , we have

$$w = Iv, \quad \beta_k^{(i)} = (E_i R_i v)(x_k) - v(x_k) \quad (16)$$

for all  $k \in \mathcal{I}_i^\Gamma$ ,  $i = 1, 2$ .

*Proof* From  $V_i^\Gamma \subset R_i V$ ,  $V \subset R_1 V \oplus R_2 V$  and (8) it follows that  $V + V_1^\Gamma + V_2^\Gamma \subset V^\Gamma$  holds. We now show that each  $v \in V^\Gamma$  can be decomposed as

$$v = w + v_1^\Gamma + v_2^\Gamma \quad \text{with } w \in V, v_i^\Gamma \in V_i^\Gamma, \quad (17)$$

and that this decomposition is unique and has the form (15), (16). We use as ansatz a decomposition as in (17). Due to  $v_i^\Gamma(x_k) = 0$  for all  $x_k \in \mathcal{V}$  we obtain  $w = Iv$ . From  $v - Iv \in V^\Gamma = R_1 V \oplus R_2 V$  and  $\text{supp}(v - Iv) \subset \Omega_\Gamma$  it follows that  $v - Iv \in V_1^\Gamma + V_2^\Gamma$  and thus a decomposition as in (17) exists. We get

$$\begin{aligned} v &= R_1 v + R_2 v \\ &= R_1 Iv + R_2 Iv \\ &\quad + R_1 \left( \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k^{(1)} \phi_k \right) + R_2 \left( \sum_{k \in \mathcal{I}_2^\Gamma} \beta_k^{(2)} \phi_k \right), \end{aligned}$$

with suitable coefficients  $\beta_k^{(i)}$ . Thus

$$E_i R_i v = E_i R_i Iv + E_i R_i \left( \sum_{k \in \mathcal{I}_i^\Gamma} \beta_k^{(i)} \phi_k \right), \quad i = 1, 2,$$

must hold. Using (13) this yields

$$E_i R_i v = (Iv)|_{\Omega_i^e} + \sum_{k \in \mathcal{I}_i^\Gamma} \beta_k^{(i)} (\phi_k)|_{\Omega_i^e}, \quad i = 1, 2.$$

Substitution of a vertex  $x_k$  with  $k \in \mathcal{I}_i^\Gamma$  yields  $\beta_k^{(i)} = (E_i R_i v)(x_k) - (Iv)|_{\Omega_i^e}(x_k) = (E_i R_i v)(x_k) - v(x_k)$ , which completes the proof.

We derive an optimal approximation error bound for the XFEM space  $V^\Gamma$ .

**Theorem 3** *For integers  $0 \leq l < m \leq 2$  the following holds*

$$\inf_{v \in V^\Gamma} \|u - v\|_{l, \Omega_1 \cup \Omega_2} \leq c h^{m-l} \|u\|_{m, \Omega_1 \cup \Omega_2}, \quad (18)$$

for all  $u \in H^m(\Omega_1 \cup \Omega_2)$ .

*Proof* We use extension operators  $\mathcal{E}_i^m : H^m(\Omega_i) \rightarrow H^m(\Omega)$ ,  $i = 1, 2$ , with  $(\mathcal{E}_i^m w)|_{\Omega_i} = w$  and  $\|\mathcal{E}_i^m w\|_m \leq c \|w\|_{m, \Omega_i}$ , cf. [12]. For  $m = 1, 2$ , let  $Q^m : H^m(\Omega) \rightarrow V$  be a (quasi-)interpolation operator such that  $\|w - Q^m w\|_l \leq c h^{m-l} \|w\|_m$  for all  $w \in H^m(\Omega)$ ,  $0 \leq l < m \leq 2$  (for example, Clement quasi-interpolation). Let  $m \in \{1, 2\}$  and  $u \in H^m(\Omega_1 \cup \Omega_2)$  be given. Define  $v^* \in V^\Gamma$  by

$$v^* = R_1 Q^m \mathcal{E}_1^m R_1 u + R_2 Q^m \mathcal{E}_2^m R_2 u. \quad (19)$$

For this approximation we obtain

$$\begin{aligned} &\|u - v^*\|_{l, \Omega_1 \cup \Omega_2}^2 \\ &= \sum_{i=1}^2 \|u - v^*\|_{l, \Omega_i}^2 = \sum_{i=1}^2 \|u - Q^m \mathcal{E}_i^m R_i u\|_{l, \Omega_i}^2 \\ &= \sum_{i=1}^2 \|\mathcal{E}_i^m R_i u - Q^m \mathcal{E}_i^m R_i u\|_{l, \Omega_i}^2 \\ &\leq \sum_{i=1}^2 \|\mathcal{E}_i^m R_i u - Q^m \mathcal{E}_i^m R_i u\|_i^2 \\ &\leq c h^{2(m-l)} \sum_{i=1}^2 \|\mathcal{E}_i^m R_i u\|_m^2 \leq c h^{2(m-l)} \sum_{i=1}^2 \|R_i u\|_{m, \Omega_i}^2 \\ &= c h^{2(m-l)} \|u\|_{m, \Omega_1 \cup \Omega_2}^2, \end{aligned}$$

which proves the result.

In [8] related approximation results for the XFEM space are derived in a mesh dependent norm in which, for  $u \in H^1(\Omega) \cap H^2(\Omega_1 \cup \Omega_2)$ , a control of the error across  $\Gamma$  is included.

### 3 A modified XFEM space

In general there are basis functions  $R_i \phi_k \in V_i^\Gamma$  with very small support in the sense that

$$|\text{supp}(R_i \phi_k)| / |\text{supp}(\phi_k)| \ll 1.$$

It is clear that if functions with ‘‘very small’’ support are deleted from the space  $V_i^\Gamma$  this will not influence the approximation quality of  $V^\Gamma$  significantly. In this section we introduce and analyze a smaller space in which basis functions from  $V_i^\Gamma$  with very small support are deleted. Avoiding very small supports has advantages, for example if the contributions are dominated by rounding errors. We will explain how we chose the maximal size of these ‘‘small supports’’ in order to maintain optimal approximation properties of the resulting reduced XFEM space.

Let  $\alpha > 0$ ,  $\tilde{c} > 0$  be given parameters. Let  $\mathcal{I}_i^\gamma \subset \mathcal{I}_i^\Gamma$  be the index set such that for all  $k \in \mathcal{I}_i^\Gamma \setminus \mathcal{I}_i^\gamma$ :

$$\frac{\|\phi_k\|_{l, T \cap \Omega_i}}{\|\phi_k\|_{l, T}} \leq \tilde{c} h_T^\alpha \quad \text{for all } T \subset (\text{supp}(\phi_k) \cap \Omega_\Gamma). \quad (20)$$

**Remark 2** Note that for a function  $R_i \phi_k \in V_i^\Gamma$  ( $k \in \mathcal{I}_i^\Gamma$ ) we have  $\|R_i \phi_k\|_{l,T} = \|\phi_k\|_{l,T \cap \Omega_i}$  for all  $T \in \mathcal{T}_h$ . Furthermore, because  $\|\phi_k\|_{l,T} \sim ch_T^{\frac{1}{2}-l}$ , for  $l = 0, 1$ , the condition (20) can be replaced by the following one:

$$\|\phi_k\|_{l,T \cap \Omega_i} \leq \hat{c} h_T^{\alpha+1-\frac{1}{2}-l} \quad (21)$$

for all  $T \subset (\text{supp}(\phi_k) \cap \Omega_\Gamma)$ .

The constant  $\hat{c}$  may differ from  $\tilde{c}$  in (20).

We define the reduced spaces  $V_i^\gamma \subset V_i^\Gamma$  by

$$V_i^\gamma := \text{span}\{R_i \phi_k \mid k \in \mathcal{I}_i^\gamma\}, \quad i = 1, 2,$$

and a modified XFEM space  $\tilde{V}^\Gamma := V \oplus V_1^\gamma \oplus V_2^\gamma$ . For this space the approximation property in Theorem 4 holds. In the analysis we use a global inverse inequality and therefore we need the additional assumption that the hierarchy of triangulations is quasi-uniform.

**Theorem 4** *We assume  $\{\mathcal{T}_h\}_{h>0}$  to be quasi-uniform. For  $0 \leq l < m \leq 2$  the following holds:*

$$\inf_{v \in \tilde{V}^\Gamma} \|u - v\|_{l, \Omega_1 \cup \Omega_2} \leq c(h^{m-l} + h^{\alpha-l}) \|u\|_{m, \Omega_1 \cup \Omega_2}$$

for all  $u \in H^m(\Omega_1 \cup \Omega_2)$ .

*Proof* Take  $u \in H^m(\Omega_1 \cup \Omega_2)$  and  $0 \leq l < m \leq 2$ . Let

$$v^* = R_1 Q^m \mathcal{E}_1^m R_1 u + R_2 Q^m \mathcal{E}_2^m R_2 u \in V^\Gamma \quad (22)$$

be as in (19). From Theorem 2 we have the representation

$$v^* = I v^* + \sum_{i=1}^2 \sum_{k \in \mathcal{I}_i^\Gamma} \beta_k^{(i)} R_i \phi_k,$$

with  $\beta_k^{(i)}$  as in (16). Given this  $v^*$  and these coefficients, we define  $\tilde{v}^* \in \tilde{V}^\Gamma$  by

$$\tilde{v}^* = I v^* + \sum_{i=1}^2 \sum_{k \in \mathcal{I}_i^\gamma} \beta_k^{(i)} R_i \phi_k.$$

Note that

$$\begin{aligned} \inf_{v \in \tilde{V}^\Gamma} \|u - v\|_{l, \Omega_1 \cup \Omega_2} &\leq \|u - \tilde{v}^*\|_{l, \Omega_1 \cup \Omega_2} \\ &\leq \|u - v^*\|_{l, \Omega_1 \cup \Omega_2} + \|v^* - \tilde{v}^*\|_{l, \Omega_1 \cup \Omega_2} \\ &\leq ch^{m-l} \|u\|_{m, \Omega_1 \cup \Omega_2} + \|v^* - \tilde{v}^*\|_{l, \Omega_1 \cup \Omega_2}. \end{aligned} \quad (23)$$

For the last term on the right hand-side in (23) we have

$$\|v^* - \tilde{v}^*\|_{l, \Omega_1 \cup \Omega_2} \leq \sum_{i=1}^2 \left\| \sum_{k \in \mathcal{I}_i^\Gamma \setminus \mathcal{I}_i^\gamma} \beta_k^{(i)} R_i \phi_k \right\|_{l, \Omega_1 \cup \Omega_2}.$$

We take  $i = 1$  (the case  $i = 2$  can be treated in the same way):

$$\begin{aligned} &\left\| \sum_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} \beta_k^{(1)} R_1 \phi_k \right\|_{l, \Omega_1 \cup \Omega_2}^2 \\ &= \sum_{T \in \Omega_\Gamma} \left\| \sum_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} \beta_k^{(1)} R_1 \phi_k \right\|_{l, T \cap \Omega_1}^2 \\ &\leq 3 \max_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} |\beta_k^{(1)}|^2 \sum_{T \in \Omega_\Gamma} \sum_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} \|\phi_k\|_{l, T \cap \Omega_1}^2 \\ &\leq c \max_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} |\beta_k^{(1)}|^2 \sum_{T \in \Omega_\Gamma} h_T^{2\alpha+3-2l} \\ &\leq c \max_{k \in \mathcal{I}_1^\Gamma \setminus \mathcal{I}_1^\gamma} |\beta_k^{(1)}|^2 h^{2\alpha+1-2l}. \end{aligned} \quad (24)$$

In the last inequality we used that  $\sum_{T \in \Omega_\Gamma} h_T^2 \leq c \text{meas}_2(\Gamma)$  holds. For  $\beta_k^{(1)}$  we have from (16)

$$|\beta_k^{(1)}| = |(E_1 R_1 v^*)(x_k) - v^*(x_k)|,$$

and using (22) we get, for  $k \in \mathcal{I}_1^\Gamma$ ,

$$\begin{aligned} (E_1 R_1 v^*)(x_k) &= (E_1 R_1 Q^m \mathcal{E}_1^m R_1 u)(x_k) \\ &= (Q^m \mathcal{E}_1^m R_1 u)(x_k), \\ v^*(x_k) &= (R_2 v^*)(x_k) = (Q^m \mathcal{E}_2^m R_2 u)(x_k). \end{aligned}$$

Hence, using  $\|w\|_{L^\infty} \leq ch^{-\frac{1}{2}} \|w\|_1$  for  $w \in V$  (cf. [2]), we obtain

$$\begin{aligned} |\beta_k^{(1)}|^2 &\leq \max_{1 \leq k \leq n} |Q^m (\mathcal{E}_1^m R_1 u - \mathcal{E}_2^m R_2 u)(x_k)|^2 \\ &= \|Q^m (\mathcal{E}_1^m R_1 u - \mathcal{E}_2^m R_2 u)\|_{L^\infty}^2 \\ &\leq ch^{-1} \|Q^m (\mathcal{E}_1^m R_1 u - \mathcal{E}_2^m R_2 u)\|_1^2 \\ &\leq ch^{-1} (\|\mathcal{E}_1^m R_1 u\|_1^2 + \|\mathcal{E}_2^m R_2 u\|_1^2) \\ &\leq ch^{-1} (\|\mathcal{E}_1^m R_1 u\|_m^2 + \|\mathcal{E}_2^m R_2 u\|_m^2) \\ &\leq ch^{-1} (\|R_1 u\|_{m, \Omega_1}^2 + \|R_2 u\|_{m, \Omega_2}^2) \\ &= ch^{-1} \|u\|_{m, \Omega_1 \cup \Omega_2}^2. \end{aligned}$$

Using this in (24) we obtain

$$\|v^* - \tilde{v}^*\|_{l, \Omega_1 \cup \Omega_2}^2 \leq ch^{2(\alpha-l)} \|u\|_{m, \Omega_1 \cup \Omega_2}^2,$$

which combined with (23) completes the proof.

From this theorem we conclude that the order of approximation of the modified XFEM space  $\tilde{V}^\Gamma$  is the same as that of  $V^\Gamma$  if we take  $\alpha = m$  in the criterion (20) (or (21)). In practice one has to choose the constant  $\tilde{c}$  in (20) (or  $\hat{c}$  in (21)). In our applications the modified XFEM space is used for the discretization of the pressure  $p \in H^1(\Omega_1 \cup \Omega_2)$ , i.e., we take  $\alpha = m = 1$ .

#### 4 $L^2$ -Stability of a basis in the XFEM space

In this section we analyze the stability of the basis

$$\{\phi_k\}_{1 \leq k \leq n} \cup \{R_1 \phi_k\}_{k \in \mathcal{I}_1^T} \cup \{R_2 \phi_k\}_{k \in \mathcal{I}_2^T}$$

in the space  $V^T$ . We prove that the diagonally scaled mass matrix is uniformly (w.r.t.  $h$ ) well-conditioned. This holds independent of the size and the shape of the support of the basis functions  $R_i \phi_k$ . This immediately implies a similar result for the reduced XFEM space  $\tilde{V}^T$ .

We start with an elementary lemma.

**Lemma 1** *Let  $T$  be a nondegenerated tetrahedron with vertices  $A, B, C, D$ . Let  $v, w : \mathbb{R}^3 \rightarrow \mathbb{R}$  be linear functions with  $v(A) = \alpha \neq 0$ ,  $v(B) = \beta$ ,  $v(C) = v(D) = 0$  and  $w(D) = \delta \neq 0$ ,  $w(A) = w(B) = w(C) = 0$ . Then the following holds:*

$$\begin{aligned} |(v, w)_{0,T}| &= \frac{1}{\sqrt{2}} \left( \frac{(\alpha + \beta)^2}{(\alpha + \beta)^2 + \alpha^2 + \beta^2} \right)^{\frac{1}{2}} \|v\|_{0,T} \|w\|_{0,T} \\ &\leq \frac{1}{\sqrt{2}} \|v\|_{0,T} \|w\|_{0,T}. \end{aligned} \quad (25)$$

*Proof* By a scaling argument we can assume  $\delta = 1$ . Let  $M_i$ ,  $i = 1, \dots, 6$  be the midpoints of the edges of  $T$ . We use the quadrature formula  $Q_T(f) = |T| \left( \frac{1}{20}(f(A) + f(B) + f(C) + f(D)) + \frac{1}{5} \sum_{i=1}^6 f(M_i) \right)$ , which is exact for all  $f \in \mathcal{P}_2$ . A simple computation using this quadrature formula yields

$$\begin{aligned} (v, w)_{0,T} &= |T| \frac{1}{20} (\alpha + \beta) \\ \|v\|_{0,T}^2 &= |T| \frac{1}{10} (\alpha^2 + \beta^2 + \alpha\beta) \\ \|w\|_{0,T}^2 &= |T| \frac{1}{10}. \end{aligned}$$

Hence

$$\frac{|(v, w)_{0,T}|}{\|v\|_{0,T} \|w\|_{0,T}} = \frac{1}{\sqrt{2}} \frac{|\alpha + \beta|}{\sqrt{(\alpha + \beta)^2 + \alpha^2 + \beta^2}},$$

which proves the desired result.

This lemma shows that a strengthened Cauchy-Schwarz inequality holds. Note that if  $\beta = 0$  the constant can be improved from  $\frac{1}{\sqrt{2}}$  to  $\frac{1}{2}$ .

We also derive a strengthened Cauchy-Schwarz inequality between the spaces  $V$  and  $V_1^T \oplus V_2^T$ . For this we introduce the following technical assumption. For each vertex  $x \in \mathcal{V}$  let  $\omega_x$  be the set of all tetrahedra that have  $x$  as a vertex. Define  $\Omega_R = \Omega \setminus \Omega_\Gamma$ . Assume that

$$\omega_x \cap \Omega_R \neq \emptyset \quad \text{for all } x \in \mathcal{V}. \quad (26)$$

For  $h$  sufficiently small this assumption is satisfied.

**Lemma 2** *Assume that (26) holds. There exists a constant  $c_{CS} < 1$  independent of  $h$  such that*

$$(v, w)_0 \leq c_{CS} \|v\|_0 \|w\|_0 \quad \text{for all } v \in V, w \in V_1^T \oplus V_2^T.$$

*Proof* We use the notation  $W = V_1^T \oplus V_2^T$ . Let  $P_W : L^2(\Omega) \rightarrow W$  be the  $L^2$ -orthogonal projection on  $W$ . Let  $\mathcal{V}(T)$  denote the set of vertices of  $T$ . Transformation to a unit tetrahedron yields the norm equivalence

$$c_1 \|v\|_{0,T}^2 \leq |T| \sum_{x \in \mathcal{V}(T)} v(x)^2 \leq c_2 \|v\|_{0,T}^2 \quad (27)$$

for all  $T \in \mathcal{T}_h$ ,  $v \in V$ , with constants  $c_1 > 0$  and  $c_2$  independent of  $h$ . Due to (26) we have that for each  $x \in \mathcal{V}(T)$  there exists a tetrahedron  $\hat{T} \in \omega_x \cap \Omega_R$  with  $x \in \mathcal{V}(\hat{T})$ . Using this we obtain for  $v \in V$  and  $T \in \Omega_\Gamma$ :

$$\begin{aligned} \|v\|_{0,T}^2 &\leq c |T| \sum_{x \in \mathcal{V}(T)} v(x)^2 \\ &\leq c \sum_{x \in \mathcal{V}(T)} \sum_{\hat{T} \in \omega_x \cap \Omega_R} |\hat{T}| \sum_{y \in \mathcal{V}(\hat{T})} v(y)^2 \\ &\leq c \sum_{x \in \mathcal{V}(T)} \|v\|_{0,\omega_x \cap \Omega_R}^2. \end{aligned}$$

Hence,

$$\|v\|_{0,\Omega_\Gamma}^2 = \sum_{T \in \Omega_\Gamma} \|v\|_{0,T}^2 \leq c \|v\|_{0,\Omega_R}^2, \quad v \in V,$$

holds with a constant  $c$  independent of  $h$ . This yields  $\|v\|_0^2 = \|v\|_{0,\Omega_\Gamma}^2 + \|v\|_{0,\Omega_R}^2 \leq c \|v\|_{0,\Omega_R}^2$  with  $c$  independent of  $h$ . Using this and  $(P_W v)|_{\Omega_R} = 0$  we get, for  $v \in V$ ,

$$\|v - P_W v\|_0 \geq \|v - P_W v\|_{0,\Omega_R} = \|v\|_{0,\Omega_R} \geq \hat{c} \|v\|_0,$$

with a constant  $\hat{c} > 0$  independent of  $h$ . Thus we get

$$\|P_W v\|_0^2 = \|v\|_0^2 - \|v - P_W v\|_0^2 \leq (1 - \hat{c}^2) \|v\|_0^2 =: c_{CS}^2 \|v\|_0^2$$

for all  $v \in V$ , and, for  $v \in V$ ,  $w \in W$ ,

$$\begin{aligned} (v, w)_0 &= (v, P_W w)_0 = (P_W v, w)_0 \leq \|P_W v\|_0 \|w\|_0 \\ &\leq c_{CS} \|v\|_0 \|w\|_0, \end{aligned}$$

which completes the proof.

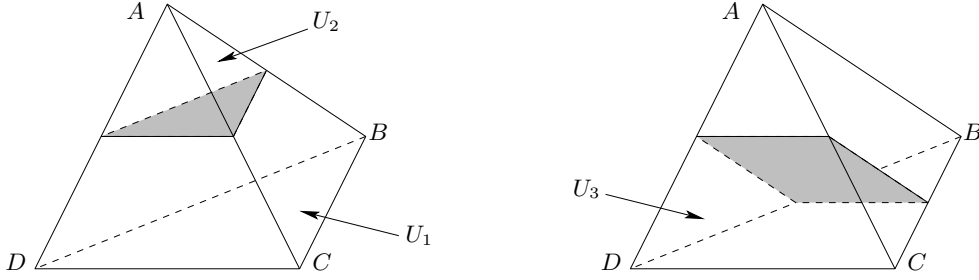
The spaces  $V_1^T$  and  $V_2^T$  are (due to disjoint supports of functions from these spaces)  $L^2$ -orthogonal. Thus we conclude that in the decomposition

$$V^T = V \oplus V_1^T \oplus V_2^T$$

we have a strengthened Cauchy-Schwarz inequality between  $V$  and  $V_1^T \oplus V_2^T$  and even orthogonality between  $V_1^T$  and  $V_2^T$ .

For  $v = w + w_1 + w_2 \in V^T$ , with  $w \in V$ ,  $w_i \in V_i^T$ , we have

$$\|v\|_0^2 \leq 2(\|w\|_0^2 + \|w_1 + w_2\|_0^2) = 2(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2)$$



**Fig. 3** Three different cases for  $U_j = T \cap \Omega_1$ ,  $j = 1, 2, 3$ .

and

$$\begin{aligned} \|v\|_0^2 &= \|w\|_0^2 + \|w_1 + w_2\|_0^2 + 2(w, w_1 + w_2)_0 \\ &\geq \|w\|_0^2 + \|w_1 + w_2\|_0^2 - 2c_{CS}\|w\|_0\|w_1 + w_2\|_0 \\ &\geq (1 - c_{CS})(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2). \end{aligned}$$

Hence we obtain

$$\begin{aligned} (1 - c_{CS})(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2) \\ \leq \|v\|_0^2 \leq 2(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2). \end{aligned} \quad (28)$$

We now turn to the conditioning of the mass matrix. A function  $v \in V^T$  is represented in the basis  $\{\phi_k\}_{1 \leq k \leq n} \cup \{R_1\phi_k\}_{k \in \mathcal{I}_1^T} \cup \{R_2\phi_k\}_{k \in \mathcal{I}_2^T}$  as

$$v = \sum_{k=1}^n \alpha_k \phi_k + \sum_{i=1}^2 \sum_{k \in \mathcal{I}_i^T} \beta_k^{(i)} R_i \phi_k =: w + w_1 + w_2, \quad (29)$$

where  $w \in V$ ,  $w_i \in V_i^T$ ,  $i = 1, 2$ . It is well-known (cf. also (27)) that for  $w = \sum_{k=1}^n \alpha_k \phi_k$  we have

$$c_1 \sum_{k=1}^n \alpha_k^2 \|\phi_k\|_0^2 \leq \|w\|_0^2 \leq c_2 \sum_{k=1}^n \alpha_k^2 \|\phi_k\|_0^2, \quad (30)$$

with constants  $c_1 > 0$  and  $c_2$  independent of  $h$ , i.e., the basis  $\{\phi_k\}_{1 \leq k \leq n}$  is uniformly in  $h$  well-conditioned (w.r.t.  $\|\cdot\|_0$ ). We prove a similar result for the basis  $\{R_i\phi_k\}_{k \in \mathcal{I}_i^T}$  of  $V_i^T$ .

**Lemma 3** For  $w_i = \sum_{k \in \mathcal{I}_i^T} \beta_k^{(i)} R_i \phi_k$ ,  $i = 1, 2$ , the following holds:

$$\begin{aligned} \frac{\sqrt{2}-1}{2\sqrt{2}} \sum_{k \in \mathcal{I}_1^T} (\beta_k^{(1)})^2 \|R_1 \phi_k\|_0^2 \\ \leq \|w_1\|_0^2 \leq 3 \sum_{k \in \mathcal{I}_1^T} (\beta_k^{(1)})^2 \|R_1 \phi_k\|_0^2. \end{aligned} \quad (31)$$

*Proof* We take  $i = 1$  and write  $w_1 = \sum_{k \in \mathcal{I}_1^T} \beta_k R_1 \phi_k$ . For each  $T \in \Omega_T$  there are at most 3  $k$ -values in  $\mathcal{I}_1^T$

with  $(R_1\phi_k)|_T \neq 0$  and thus

$$\begin{aligned} \|w_1\|_0^2 &= \sum_{T \in \Omega_T} \left\| \sum_{k \in \mathcal{I}_1^T} \beta_k R_1 \phi_k \right\|_{0,T}^2 \\ &\leq \sum_{T \in \Omega_T} \left( \sum_{k \in \mathcal{I}_1^T} |\beta_k| \|R_1 \phi_k\|_{0,T} \right)^2 \\ &\leq 3 \sum_{T \in \Omega_T} \sum_{k \in \mathcal{I}_1^T} |\beta_k|^2 \|R_1 \phi_k\|_{0,T}^2 \\ &= 3 \sum_{k \in \mathcal{I}_1^T} |\beta_k|^2 \|R_1 \phi_k\|_0^2, \end{aligned}$$

which proves the upper bound in (31). We now prove the lower bound. For a given  $T \in \Omega_T$  we consider

$$\left\| \sum_{k \in \mathcal{I}_1^T} \beta_k R_1 \phi_k \right\|_{0,T} = \left\| \sum_{k \in \mathcal{I}_1^T} \beta_k \phi_k \right\|_{0, T \cap \Omega_1}.$$

For  $T \cap \Omega_1$  there are 3 different cases, as indicated in Fig. 3, namely  $T \cap \Omega_1 = U_j$ ,  $j = 1, 2, 3$ .

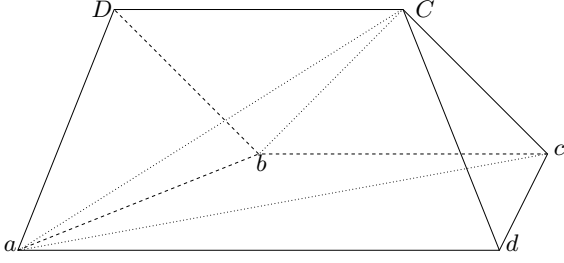
For  $T \cap \Omega_1 = U_1$  there is only *one*  $k \in \mathcal{I}_1^T$  with  $(\phi_k)|_{T \cap \Omega_1} \neq 0$  (namely the one corresponding to vertex  $A$ ). Hence in this case we have

$$\left\| \sum_{k \in \mathcal{I}_1^T} \beta_k \phi_k \right\|_{0, T \cap \Omega_1}^2 = \sum_{k \in \mathcal{I}_1^T} \beta_k^2 \|\phi_k\|_{0, T \cap \Omega_1}^2.$$

We now treat  $T \cap \Omega_1 = U_2$ . Note that  $U_2$  is a tetrahedron. There are three  $k$ -values in  $\mathcal{I}_1^T$  with  $(\phi_k)|_{T \cap \Omega_1} \neq 0$ , say  $k_B, k_C, k_D$ , corresponding to the nodal basis functions  $\phi_k$  at the vertices  $B, C, D$ , respectively. Using Lemma 1 we get

$$\begin{aligned} \left\| \sum_{k \in \mathcal{I}_1^T} \beta_k \phi_k \right\|_{0, T \cap \Omega_1}^2 \\ = \|\beta_{k_B} \phi_{k_B} + (\beta_{k_C} \phi_{k_C} + \beta_{k_D} \phi_{k_D})\|_{0, T \cap \Omega_1}^2 \\ \geq (1 - 1/\sqrt{2})(\beta_{k_B}^2 \|\phi_{k_B}\|_{0, T \cap \Omega_1}^2 \\ + \|\beta_{k_C} \phi_{k_C} + \beta_{k_D} \phi_{k_D}\|_{0, T \cap \Omega_1}^2), \end{aligned}$$





**Fig. 4** Pentahedron  $U_3$ , subdivided in three tetrahedra  $T_1, T_2, T_3$ .

and, again with Lemma 1 (with  $\beta = 0$ ), we obtain

$$\begin{aligned} & \|\beta_{k_C} \phi_{k_C} + \beta_{k_D} \phi_{k_D}\|_{0, T \cap \Omega_1}^2 \\ & \geq \frac{1}{2} (\beta_{k_C}^2 \|\phi_{k_C}\|_{0, T \cap \Omega_1}^2 + \beta_{k_D}^2 \|\phi_{k_D}\|_{0, T \cap \Omega_1}^2), \end{aligned}$$

and thus

$$\left\| \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k \phi_k \right\|_{0, T \cap \Omega_1}^2 \geq \frac{\sqrt{2}-1}{2\sqrt{2}} \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k^2 \|\phi_k\|_{0, T \cap \Omega_1}^2.$$

Finally we consider  $T \cap \Omega_1 = U_3$ . In this case there are two  $k$ -values in  $\mathcal{I}_1^\Gamma$  with  $(\phi_k)|_{T \cap \Omega_1} \neq 0$ , say  $k_A, k_B$ , corresponding to the nodal basis functions  $\phi_k$  at the vertices  $A, B$ , respectively. The pentahedron  $U_3$  has the form as indicated in Fig. 4 and can be subdivided into 3 tetrahedra  $T_1 = adcC$ ,  $T_2 = acCb$ ,  $T_3 = aCbD$  (see Fig. 4).

We consider  $T_1$ . The basis function  $\phi_{k_A}$  has nonzero values at the vertices  $a$  and  $d$  of  $T_1$  and is zero at the vertices  $c$  and  $C$ . The basis function  $\phi_{k_B}$  has a nonzero value at vertex  $c$  and zero values at  $a, d$  and  $C$ . Thus Lemma 1 can be applied and results in

$$(\phi_{k_A}, \phi_{k_B})_{0, T_1} \leq \frac{1}{\sqrt{2}} \|\phi_{k_A}\|_{0, T_1} \|\phi_{k_B}\|_{0, T_1}.$$

It can be checked that the same argument can be applied to  $T_2$  and  $T_3$ . Using this we get

$$\begin{aligned} & \left\| \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k \phi_k \right\|_{0, T \cap \Omega_1}^2 = \|\beta_{k_A} \phi_{k_A} + \beta_{k_B} \phi_{k_B}\|_{0, T \cap \Omega_1}^2 \\ & = \sum_{j=1}^3 \|\beta_{k_A} \phi_{k_A} + \beta_{k_B} \phi_{k_B}\|_{0, T_j}^2 \\ & \geq (1 - 1/\sqrt{2}) \sum_{j=1}^3 (\beta_{k_A}^2 \|\phi_{k_A}\|_{0, T_j}^2 + \beta_{k_B}^2 \|\phi_{k_B}\|_{0, T_j}^2) \\ & = (1 - 1/\sqrt{2}) (\beta_{k_A}^2 \|\phi_{k_A}\|_{0, T \cap \Omega_1}^2 + \beta_{k_B}^2 \|\phi_{k_B}\|_{0, T \cap \Omega_1}^2) \\ & = (1 - 1/\sqrt{2}) \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k^2 \|\phi_k\|_{0, T \cap \Omega_1}^2. \end{aligned}$$

We conclude that for all three cases  $T \cap \Omega_1 = U_j$ ,  $j = 1, 2, 3$ , we have

$$\left\| \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k R_1 \phi_k \right\|_{0, T}^2 \geq \frac{\sqrt{2}-1}{2\sqrt{2}} \sum_{k \in \mathcal{I}_1^\Gamma} \beta_k^2 \|R_1 \phi_k\|_{0, T}^2.$$

Summing over  $T \in \Omega_\Gamma$  we obtain the lower bound in (31).

Using the norm equivalences in (28), (30) and (31) we derive a spectral result for the mass matrix using standard arguments. Let  $m := n + |\mathcal{I}_1^\Gamma| + |\mathcal{I}_2^\Gamma|$  be the dimension of  $V^\Gamma$  and  $P : \mathbb{R}^m \rightarrow V^\Gamma$  the isomorphism defined by (29):

$$Pz = P(\boldsymbol{\alpha}, \boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}) = v.$$

The mass matrix  $M \in \mathbb{R}^{m \times m}$  is given by

$$\langle Mz, z \rangle = (Pz, Pz)_0 \quad \text{for all } z \in \mathbb{R}^m.$$

Here  $\langle \cdot, \cdot \rangle$  denotes the Euclidean scalar product. Define  $\text{diag}(M) =: D_M$  with

$$D_M = \begin{pmatrix} D & \emptyset \\ \emptyset & D_2 \end{pmatrix}, \quad D_{k,k} = \|\phi_k\|_0^2, \quad 1 \leq k \leq n, \\ (D_i)_{k,k} = \|R_i \phi_k\|_0^2, \quad k \in \mathcal{I}_i^\Gamma.$$

**Theorem 5** *There are constants  $c_1 > 0$  and  $c_2$  independent of  $h$  such that*

$$c_1 \langle D_M z, z \rangle \leq \langle Mz, z \rangle \leq c_2 \langle D_M z, z \rangle \quad \text{for all } z \in \mathbb{R}^m.$$

*Proof* From (28), (30) and (31) we get

$$\begin{aligned} \langle Mz, z \rangle & = \|v\|_0^2 \leq 2(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2) \\ & \leq 2 \left( c_2 \sum_{k=1}^n \alpha_k^2 \|\phi_k\|_0^2 + 3 \sum_{k \in \mathcal{I}_1^\Gamma} (\beta_k^{(1)})^2 \|R_1 \phi_k\|_0^2 \right. \\ & \quad \left. + 3 \sum_{k \in \mathcal{I}_2^\Gamma} (\beta_k^{(2)})^2 \|R_2 \phi_k\|_0^2 \right) \\ & \leq c(\langle D\boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle + \langle D_1 \boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(1)} \rangle + \langle D_2 \boldsymbol{\beta}^{(2)}, \boldsymbol{\beta}^{(2)} \rangle) \\ & = c \langle D_M z, z \rangle, \end{aligned}$$

with a constant  $c$  independent of  $h$ . Similarly, due to

$$\langle Mz, z \rangle = \|v\|_0^2 \geq (1 - cCS)(\|w\|_0^2 + \|w_1\|_0^2 + \|w_2\|_0^2)$$

and using the lower bounds in (30) and (31) we obtain  $\langle Mz, z \rangle \geq c \langle D_M z, z \rangle$  with a constant  $c > 0$  independent of  $h$ .

The result in this theorem proves that the matrix  $D_M^{-1}M$  has a spectral condition number that is uniformly (w.r.t.  $h$ ) bounded. Note that the constants in the spectral condition number bounds are also independent of the supports of the basis functions  $R_i\phi_k$ ,  $k \in \mathcal{I}_i^F$ . In other words, a simple scaling is sufficient to control the stability (in  $L^2$ ) of the basis functions with “very small” supports. Furthermore, we note that in the analysis we did *not* assume quasi-uniformity of the family of triangulations.

**Corollary 1** Let  $\tilde{V}^F = V \oplus V_1^\gamma \oplus V_2^\gamma$  be the modified XFEM space introduced in section 3. Let  $\tilde{M}$  be the mass matrix w.r.t. the basis  $\{\phi_k\}_{1 \leq k \leq n} \cup \{R_1\phi_k\}_{k \in \mathcal{I}_1^\gamma} \cup \{R_2\phi_k\}_{k \in \mathcal{I}_2^\gamma}$  in this space and  $D_{\tilde{M}} = \text{diag}(\tilde{M})$ . From the analysis it trivially follows that the matrix  $D_{\tilde{M}}^{-1}\tilde{M}$  has a uniformly (w.r.t.  $h$  and size of  $T \cap \Omega_i$ ) bounded spectral condition number.

## 5 Numerical experiments

In this section we present results of numerical experiments. In section 5.1 we compute approximation errors for a given piecewise smooth function using the approximation spaces  $V$ ,  $V^F$  and  $\tilde{V}^F$ . The behavior of these approximation errors confirms the results of the theoretical analysis in the sections 2 and 3. In section 5.2 we consider a very simple Stokes two-phase flow problem (static bubble) to demonstrate the impact of an improved pressure space. In section 5.3 for the XFEM space  $V_l^F$  we present results for the spectral condition number of the scaled mass matrix. Finally, we briefly address the issue of LBB-stability of the  $(P_2)^3 - \tilde{V}_h^F$  pair by means of a numerical experiment. In all our experiments we use the cube  $\Omega = (-1, 1)^3$ .

### 5.1 Approximation of a function that is discontinuous across a planar interface

We take a planar interface  $\Gamma = \{(x, y, z) \in \Omega \mid y + z = 0.05\}$  and  $\Omega_1 = \{(x, y, z) \in \Omega \mid y + z < 0.05\}$ ,  $\Omega_2 = \Omega \setminus \Omega_1$ . Let  $u$  be given by

$$u = \begin{cases} x^2 + y^2 + z^2 & \text{in } \Omega_1 \\ 3x^2 + y^2 + 2z^2 + 2 & \text{in } \Omega_2. \end{cases}$$

We use a *uniform* triangulation of  $\Omega$  with tetrahedra, resulting in a family  $\{\mathcal{T}_{h_l}\}_{l \geq 0}$  with mesh size parameter  $h = h_l = 2^{-l}$ ,  $l = 0, 1, 2, \dots$ . The interface  $\Gamma$  and the triangulations are such that  $\Gamma$  is not aligned with the triangulation. Let  $V_l$  be the space of continuous piecewise linear functions on  $\mathcal{T}_{h_l}$  and  $V_l^F$ ,  $\tilde{V}_l^F$  the corresponding XFEM and modified XFEM spaces, respectively. For the space  $\tilde{V}_l^F$  the parameter  $\hat{c}$  in the criterion (21) has to be chosen. Below we consider different values for  $\hat{c}$ . We

$l$	$W_l = V_l$	order	$W_l = V_l^F$	order
0	2.14e+0	-	5.14e-1	-
1	1.60e+0	0.42	1.44e-1	1.83
2	1.20e+0	0.41	3.71e-2	1.96
3	8.88e-1	0.43	9.37e-3	1.99
4	6.27e-1	0.50	2.35e-3	1.99
5	4.52e-1	0.47	5.89e-4	2.00

**Table 1** Approximation errors  $e_l$  and numerical order of convergence for  $V_l$ ,  $V_l^F$ .

$l$	$\hat{c} = 10$	order	$\hat{c} = 1$	order	$\hat{c} = 0.1$	order
0	2.14e+0	-	2.14e+0	-	1.44e+0	-
1	1.60e+0	0.42	1.60e+0	0.42	1.77e-1	3.03
2	1.20e+0	0.41	2.69e-1	2.57	4.01e-2	2.14
3	8.88e-1	0.43	4.72e-2	2.51	9.37e-3	2.10
4	1.37e-2	6.01	8.98e-3	2.39	2.35e-3	1.99
5	2.60e-3	2.40	5.89e-4	3.93	5.89e-4	2.00

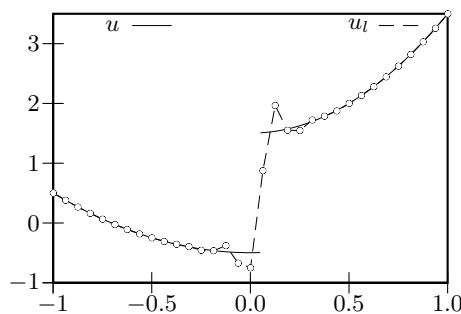
**Table 2** Approximation errors  $e_l$  and numerical order of convergence for  $\tilde{V}_l^F$ .

present approximation errors in the  $L^2$ -norm and therefore in (21) we take  $l = 0$ ,  $\alpha = 2$ . Note that for  $\hat{c} = 0$  we have  $\tilde{V}_l^F = V_l^F$  (all discontinuous basis functions are kept) and for fixed  $l$  and a sufficiently large  $\hat{c}$  we have  $\tilde{V}_l^F = V_l$  (all discontinuous basis functions are deleted). For  $W_l \in \{V_l, V_l^F, \tilde{V}_l^F\}$ . We compute  $u_l \in W_l$  such that

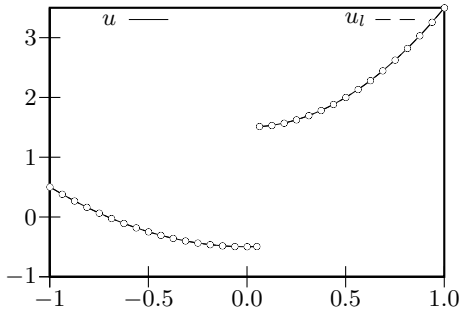
$$\|u - u_l\|_0 = \inf_{w_l \in W_l} \|u - w_l\|_0.$$

Results for the approximation error  $e_l := \|u - u_l\|_0$  are given in Table 1, Table 2. In the latter table we use the construction of the reduced space  $\tilde{V}_l^F$  based on the criterion (21) with  $\alpha = 2$  and different constants  $\hat{c} = 10, 1, 0.1$ . One-dimensional profiles of  $u_l \in V_l$  and  $u_l \in \tilde{V}_l^F$  are shown in Fig. 5 and Fig. 6.

The observed numerical order of convergence is consistent with the theoretically predicted improvement from  $p = 0.5$  to  $p = 2$ . In Table 2 we see that for  $l = 5$  the same level of accuracy can be reached if we use the reduced space  $\tilde{V}_5^F$  instead of the full extended space  $V_5^F$  and that this is not very sensitive with respect to the choice of the parameter  $\hat{c}$ .



**Fig. 5** 1D-profile of  $u_l \in V_l$  at  $x = y = 0$  for  $l = 4$ .



**Fig. 6** 1D-profile of  $u_l \in V_l^\Gamma$  at  $x = y = 0$  for  $l = 4$ .

The dimension of the space  $\tilde{V}_l^\Gamma$  depends on the value for  $\hat{c}$ . These dimensions corresponding to the spaces used in Tables 1 and 2 are given in Table 3.

$l$	$\hat{c} = \infty$	$\hat{c} = 10$	$\hat{c} = 1$	$\hat{c} = 0.1$	$\hat{c} = 0$
0	27	27	27	34	51
1	125	125	125	186	205
2	729	729	872	954	1017
3	4913	4913	5730	6001	6001
4	35937	39008	39103	40161	40161
5	274625	290878	291005	291005	291005

**Table 3** Dimension of the space  $\tilde{V}_l^\Gamma$ .

Note that for not too small  $l$  the dimension of the (modified) XFEM space is only slightly larger than that of the standard finite element space  $V_l$ .

## 5.2 A static bubble problem

In this section we consider a Stokes problem of the form (4) as described in Example 1. We take a uniform initial triangulation  $\mathcal{T}_{h_1}$  where the vertices form a  $5 \times 5 \times 5$  lattice and apply a local refinement algorithm presented in [5]. Local refinement of the coarse mesh  $\mathcal{T}_{h_1}$  in the vicinity of  $\Gamma$  yields the gradually refined meshes  $\mathcal{T}_{h_l}$ ,  $l = 2, \dots, 5$ , with local mesh sizes  $h_\Gamma = h_l = 2^{-l}$  close to the interface.

For the discretization of the velocity  $\mathbf{u}$  we choose the standard finite element space of piecewise quadratics:

$$\mathbf{V}_h := \{\mathbf{v} \in C(\Omega)^3 \mid \mathbf{v}|_T \in \mathcal{P}_2 \text{ for all } T \in \mathcal{T}_h, \mathbf{v}|_{\partial\Omega} = 0\}.$$

We describe the approximation of  $\Gamma$  by a piecewise planar manifold  $\Gamma_h$ . In level set techniques for two-phase flows the interface  $\Gamma$  is characterized as the zero level of an (approximate) signed distance function, denoted by  $d$ . In this test problem, an exact signed distance function  $d$  is known. For the construction of  $\Gamma_h$  we approximate  $d$  by a continuous piecewise *quadratic* approximation  $d_h$  (in applications this is the solution of a discretized level

set equation). Here we take the piecewise quadratic interpolation of the known function  $d$ . Let  $\mathcal{T}'_h$  be the triangulation obtained from  $\mathcal{T}_h$  after one global regular refinement and  $I(d_h)$  the continuous piecewise *linear* function on  $\mathcal{T}'_h$  that interpolates  $d_h$  (and thus, in our example, also  $d$ ) at the vertices of all tetrahedra in  $\mathcal{T}'_h$ . The approximation of  $\Gamma$  is defined by

$$\Gamma_h = \{\mathbf{x} \in \Omega \mid I(d_h)(\mathbf{x}) = 0\}.$$

Using this approximate interface  $\Gamma_h$  (with  $\text{dist}(\Gamma, \Gamma_h) \leq ch^2$ ) the discretization of the localized surface tension force  $f_\Gamma(\mathbf{v})$  is as follows. Define

$$\tilde{\mathbf{n}}_h(\mathbf{x}) := \frac{\nabla d_h(\mathbf{x})}{\|\nabla d_h(\mathbf{x})\|}, \quad \tilde{\mathbf{P}}_h(\mathbf{x}) := \mathbf{I} - \tilde{\mathbf{n}}_h(\mathbf{x})\tilde{\mathbf{n}}_h(\mathbf{x})^T,$$

where  $\mathbf{x} \in \Gamma_h$ ,  $\mathbf{x}$  not on an edge. Let  $\mathbf{e}_i$  the  $i$ -th basis vector in  $\mathbb{R}^3$  and  $(\mathbf{v}_h)_i$  the  $i$ -th component of  $\mathbf{v}_h \in \mathbf{V}_h$ . The discrete surface tension functional is given by

$$f_{\Gamma_h}(\mathbf{v}_h) = \tau \sum_{i=1}^3 \int_{\Gamma_h} \tilde{\mathbf{P}}_h(\mathbf{x}) \mathbf{e}_i \cdot \nabla_{\Gamma_h}(\mathbf{v}_h)_i ds. \quad (32)$$

In [7] it is shown that under reasonable assumptions we have the error bound

$$\sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{|f_\Gamma(\mathbf{v}_h) - f_{\Gamma_h}(\mathbf{v}_h)|}{\|\mathbf{v}_h\|_1} \leq ch_\Gamma. \quad (33)$$

We use this discrete surface tension functional in our experiments and consider the Galerkin discretization as in (6), with  $\mathbf{g} = 0$  and  $\mu = 1$ ,  $Q_h \in \{V_h, V_h^\Gamma\}$ .

In this test case the errors in velocity and pressure are influenced by two error sources, namely the approximation error of the discontinuous pressure  $p^*$  and errors induced by the discretization of the surface force  $f_\Gamma$ , cf. (7). Note that the first term in the upper bound in (7) vanishes due to  $\mathbf{u}^* = 0$ .

We consider the effect of the improved pressure finite element space  $V_h^\Gamma$  as compared to  $V_h$ . We compute the errors

$$\|e_{\mathbf{u}}\|_m = \|\mathbf{u}^* - \mathbf{u}_h\|_m = \|\mathbf{u}_h\|_m, \quad m = 0, 1,$$

and

$$\|e_p\|_0 = \|p^* - p_h\|_0,$$

with  $\mathbf{V}_h, \Gamma_h, f_{\Gamma_h}$  as explained above and with two different choices for the pressure space, namely  $V_h$  and  $V_h^\Gamma$ . Table 4 shows the decay of the pressure  $L^2$ -norm for the two different pressure spaces. The results improve significantly if we use the space  $V_h^\Gamma$ . Results for the velocity discretization error  $\|e_{\mathbf{u}}\|_m$ ,  $m = 0, 1$  with the XFEM pressure space  $V_h^\Gamma$  are given in Table 5.

In Fig. 7 and Fig. 8 a cross section of the discrete pressure solution for the two different pressure spaces is shown.

$l$	$\ e_p\ _0, p_h \in V_h$	order	$\ e_p\ _0, p_h \in V_h^F$	order
1	1.60e + 00	—	1.71e - 01	—
2	1.09e + 00	0.55	5.77e - 02	1.56
3	8.17e - 01	0.42	1.73e - 02	1.74
4	5.66e - 01	0.53	6.28e - 03	1.46
5	4.05e - 01	0.49	2.91e - 03	1.11

**Table 4** Pressure errors for the  $P_2 - V_h$  and  $P_2 - V_h^F$  finite element pair.

$l$	$\ e_u\ _0$	order	$\ e_u\ _1$	order
1	6.48e - 03	—	1.02e - 01	—
2	1.32e - 03	2.30	3.50e - 02	1.54
3	2.53e - 04	2.38	1.32e - 02	1.40
4	4.64e - 05	2.45	4.52e - 03	1.55
5	9.31e - 06	2.32	1.62e - 03	1.48

**Table 5** Velocity errors and numerical order of convergence for the  $P_2 - V_h^F$  pair.

### 5.3 Stability issues

We consider the XFEM space  $V_l^F$  ( $l = 1, \dots, 5$ ) used in the static bubble example from section 5.2. For this space we determined the mass matrix  $M_l$ . With  $D_l := \text{diag}(M_l)$  we computed the spectral condition number of  $D_l^{-1}M_l$ , i.e.,  $\kappa(D_l^{-1}M_l) = \lambda_{\max}(D_l^{-1}M_l)/\lambda_{\min}(D_l^{-1}M_l)$ . The results are given in Table 6.

$l$	$\kappa(D_l^{-1}M_l)$
1	16.16
2	11.24
3	12.08
4	12.93
5	12.98

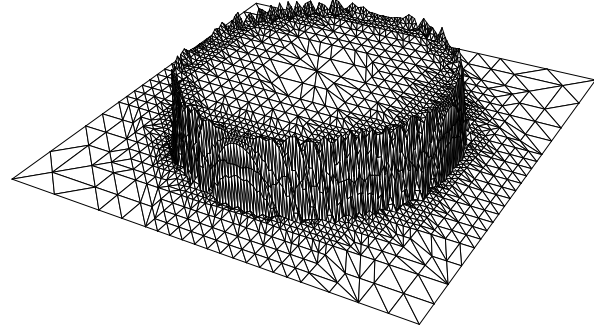
**Table 6** Spectral condition number of the scaled XFEM mass matrix.

These results clearly show the uniform boundedness of the spectral condition number of the scaled mass matrix, as proved in section 4.

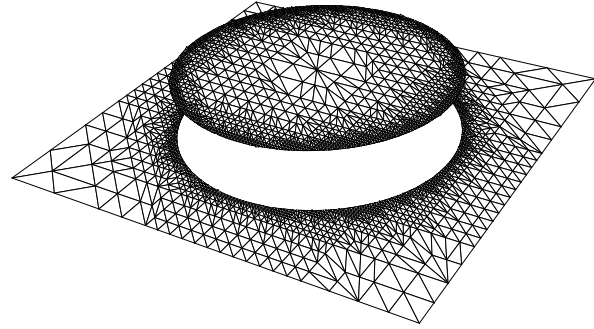
We finally briefly address a topic of current research. The standard  $\mathbf{V}_h - V_h$  Hood-Taylor pair is known to be LBB stable. An obvious question is what happens with stability if for the pressure instead of  $V_h$  we take the (larger)  $\tilde{V}_h^F$  space. We do not have a satisfactory theoretical analysis of this stability issue, yet. Here we only present results for the static bubble example from section 5.2. Let

$$K_l = \begin{pmatrix} A_l & B_l^T \\ B_l & 0 \end{pmatrix}$$

be the matrix representation of the discrete Stokes (static bubble) problem, described in the previous section, for the  $P_2 - \tilde{V}_h^F$  pair on the locally refined triangulations  $\mathcal{T}_h$  with mesh sizes  $h_\Gamma = h_l = 2^{-l}$ ,  $l = 1, \dots, 5$  close to the interface. The Schur complement matrix is given by



**Fig. 7** Finite element pressure solution  $p_h \in V_h$  on slice of  $\mathcal{T}_{h_5}^F$  at  $z = 0$ .



**Fig. 8** Finite element pressure solution  $p_h \in V_h^F$  on slice of  $\mathcal{T}_{h_5}^F$  at  $z = 0$ .

$S_l = B_l A_l^{-1} B_l^T$ . The LBB-constant for the  $\mathbf{V}_h - \tilde{V}_h^F$  pair with  $h = h_l$  is given by

$$C_{LBB}(l) = \inf_{p_h \in \tilde{V}_h^{F,*}} \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(\text{div } \mathbf{v}_h, p_h)_0}{\|\nabla \mathbf{v}_h\|_0 \|p_h\|_0},$$

where  $\tilde{V}_h^{F,*}$  contains all functions from  $\tilde{V}_h^F$  that are  $L^2$ -orthogonal to the constant. Let  $M_l$  be the mass matrix in  $\tilde{V}_h^F$  and  $m_l = \dim(\tilde{V}_h^F)$ . Define  $\mathbb{R}^{m_l,*} = \{y \in \mathbb{R}^{m_l} \mid \langle y, M_l e \rangle = 0\}$ , with  $e := (1, 1, \dots, 1)^T$ . The LBB constant can also be represented as

$$C_{LBB}^2(l) = \inf_{y \in \mathbb{R}^{m_l,*}} \frac{\langle S_l y, y \rangle}{\langle M_l y, y \rangle}, \quad (34)$$

and thus  $C_{LBB}^2(l)$  is the smallest nonzero eigenvalue of  $M_l^{-1} S_l$ . Due to the fact that  $M_l$  is uniformly spectrally

l	$\hat{c} = \infty$	$\hat{c} = 10$	$\hat{c} = 1$	$\hat{c} = 0.1$	$\hat{c} = 0.01$	$\hat{c} = 0.0001$
1	$9.53e-2$	$9.53e-2$	$9.53e-2$	$4.65e-2$	$1.43e-2$	$1.43e-2$
2	$2.53e-2$	$2.53e-2$	$2.53e-2$	$2.53e-2$	$1.53e-2$	$6.49e-3$
3	$3.22e-2$	$3.22e-2$	$3.22e-2$	$2.97e-2$	$1.07e-2$	$1.97e-4$
4	$2.58e-2$	$2.58e-2$	$2.58e-2$	$2.16e-2$	$3.17e-3$	$3.37e-5$
5	$9.17e-2$	$9.17e-2$	$5.91e-2$	$1.12e-3$	$1.60e-3$	$1.32e-5$

**Table 7** Estimates of smallest nonzero eigenvalue of preconditioned Schur complement.

equivalent to its diagonal  $D_l$  it makes sense to consider the smallest nonzero eigenvalue of  $D_l^{-\frac{1}{2}} S_l D_l^{-\frac{1}{2}}$  which is denoted by  $\lambda_{\min}^*(D_l^{-1} S_l)$ . This eigenvalue can be approximated accurately using, for example, an inverse power iteration. In each iteration of this method the linear systems with matrix  $D_l^{-\frac{1}{2}} S_l D_l^{-\frac{1}{2}}$  can be solved using a CG method. We implemented this and computed (with sufficiently high accuracy) this smallest eigenvalue for different values of the parameter  $\hat{c}$  used in the definition of  $\tilde{V}_h^T$  and for several mesh sizes. The resulting values are presented in Table 7. Note that  $\hat{c} = \infty$  corresponds to the space  $V_h$ . The rather irregular behavior in the columns in table 7 could be caused by the fact that we compute the smallest nonzero eigenvalue of  $D_l^{-1} S_l$  and not of  $M_l^{-1} S_l$ . It is clear from these results that with respect to LBB stability it is important to use the *modified* XFEM space with a not too small parameter  $\hat{c}$ . Investigation of this stability issue is a topic of current research.

**Acknowledgements** We acknowledge the help of P. Esser in performing the numerical experiments.

## References

1. Belytschko, T., Moes, N., Usui, S., Parimi, C.: Arbitrary discontinuities in finite elements. *Int. J. Num. Meth. Eng.* **50**, 993–1013 (2001)
2. Bramble, J.H., Xu, J.: Some estimates for a weighted  $l^2$ -projection. *Math. Comp.* **56**, 463–476 (1991)
3. Ganesan, S., Tobiska, L.: Finite element simulation of a droplet impinging a horizontal surface. In: *Proceedings of ALGORITMY 2005*, pp. 1–11 (2005)
4. Girault, V., Raviart, P.A.: *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin (1986)
5. Groß, S., Reusken, A.: Parallel multilevel tetrahedral grid refinement. *SIAM J. Sci. Comput.* **26**(4), 1261–1288 (2005)
6. Groß, S., Reusken, A.: An extended pressure finite element space for two-phase incompressible flows with surface tension. *J. Comp. Phys.* **224**, 40–58 (2007)
7. Groß, S., Reusken, A.: Finite element discretization error analysis of a surface tension force in two-phase incompressible flows. *SIAM J. Numer. Anal.* **45**, 1679–1700 (2007)
8. Hansbo, A., Hansbo, P.: An unfitted finite element method, based on nitsche’s method, for elliptic interface problems. *Comput. Methods Appl. Mech. Engrg.* **191**, 5537–5552 (2002)
9. Hansbo, A., Hansbo, P.: A finite element method for the simulation of strong and weak discontinuities in solid mechanics. *Comput. Methods Appl. Mech. Engrg.* **193**, 3523–3540 (2004)
10. Tornberg, A.K.: *Interface tracking methods with application to multiphase flows*. Phd thesis, Royal Institute of Technology, Department of Numerical Analysis and Computing Science, Stockholm (2000)
11. Tornberg, A.K., Engquist, B.: A finite element based level-set method for multiphase flow applications. *Comput. Visual. Sci.* **3**, 93–101 (2000)
12. Wloka, J.: *Partial Differential Equations*. Cambridge University Press, Cambridge (1992)
13. Yang, X., James, A.J., Lowengrub, J., Zheng, X., Cristini, V.: An adaptive coupled level-set/volume-of-fluid interface capturing method for unstructured triangular grids. *J. Comp. Phys.* **217**, 364–394 (2006)
14. Zheng, X., Anderson, A., Lowengrub, J., Cristini, V.: Adaptive unstructured volume remeshing II: Application to two- and three-dimensional level-set simulations of multiphase flow. *J. Comp. Phys.* **208**, 191–220 (2005)