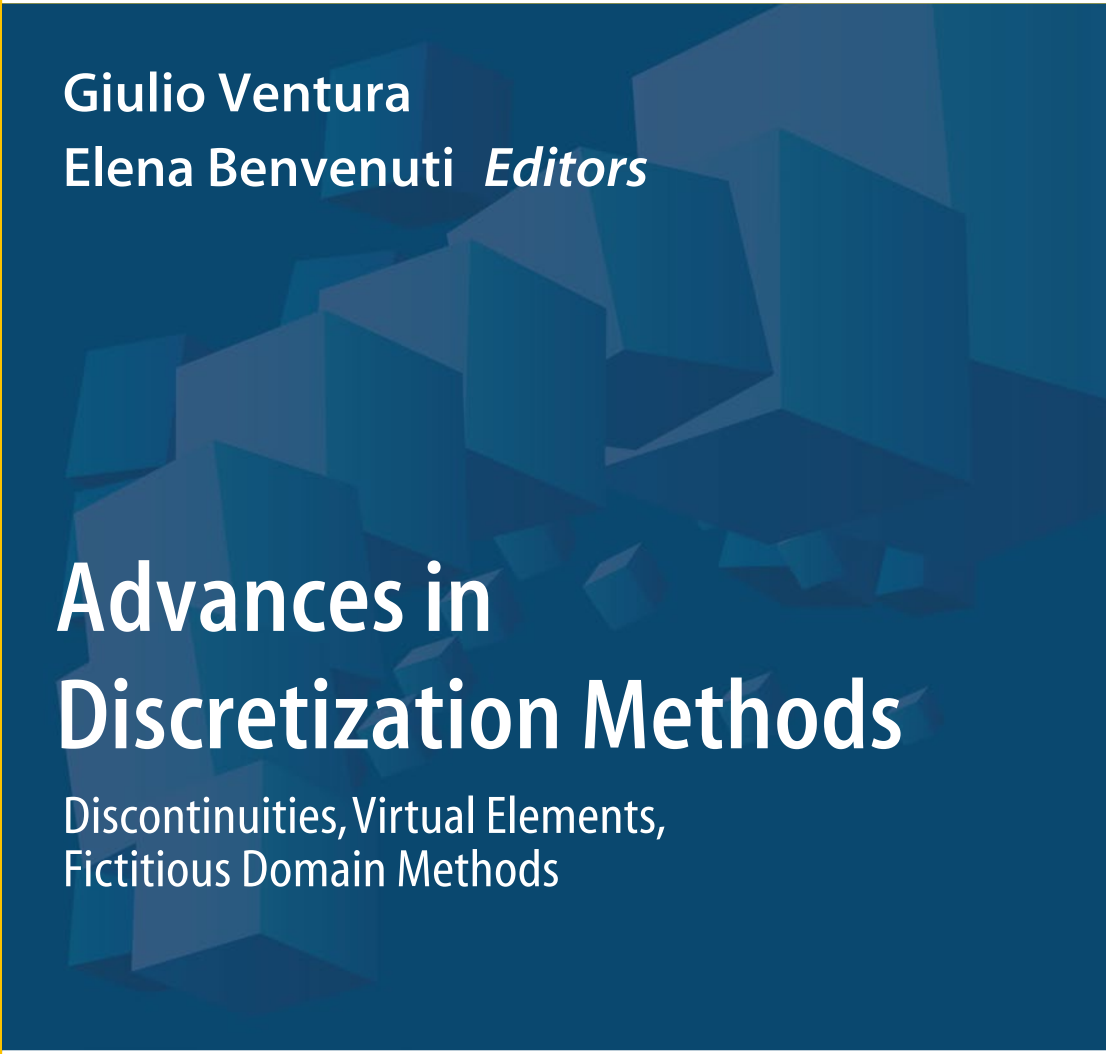
SEMA SIMAI Springer Series 12



Giulio Ventura
Elena Benvenuti *Editors*

Advances in Discretization Methods

Discontinuities, Virtual Elements, Fictitious Domain Methods

SeMA
SIMAI
SOCIETÀ ITALIANA DI MATEMATICA
APPLICATA E INDUSTRIALE
Springer

SEMA SIMAI Springer Series

Volume 12

More information about this series at http://www.springer.com/series/10532

Giulio Ventura • Elena Benvenuti
Editors

Advances in Discretization Methods

Discontinuities, Virtual Elements, Fictitious Domain Methods

Springer

Editors
Giulio Ventura
Structural, Building, Geotechnical Eng.
Politecnico di Torino
Torino, Italy

Elena Benvenuti
Department of Engineering
University of Ferrara
Ferrara, Italy

ISSN 2199-3041 ISSN 2199-305X (electronic)
SEMA SIMAI Springer Series
ISBN 978-3-319-41245-0 ISBN 978-3-319-41246-7 (eBook)
DOI 10.1007/978-3-319-41246-7

Library of Congress Control Number: 2016950209

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This volume of the SEMA SIMAI Springer Series brings together selected contributions presented at the international conference "eXtended Discretization MethodS" (X-DMS), held during September 2015 in Ferrara, Italy. The conference was one of the thematic conferences supported by the European Community in Computational Methods in Applied Sciences (ECCOMAS) and also one of the special interest conferences sponsored by the International Association for Computational Mechanics (IACM). Twelve minisymposia, more than one hundred oral presentations, and plenary lectures given by eminent personalities in the computational mechanics research field contributed to the scientific value of the event.

In gathering some of the most interesting contributions at the X-DMS 2015 Conference, the book aims to disseminate ideas and to promote discussion among researchers with an interest in the development and application of computational methods in science and technology. In line with this objective, the volume addresses some of the most advanced discretization methods for the numerical analysis of a variety of physical problems. In recent years, the efforts of the scientific community in computational mechanics have especially focused on improving both the overall computational efficiency and the versatility of the methods, including the addition of special features of the solution directly in the approximation and/or discretization space. The results of these efforts can be found in a wide range of computational methodologies, including partition of unity finite element methods (meshfree, XFEM, GFEM), virtual element and fictitious domain methods, special techniques for static and evolving interfaces, multiscale discretization, strong discontinuity approaches. The selected contributions in this volume recall the main aspects of some of these methodologies, demonstrating their potentialities and possibilities for application.

The book is organized into four parts. Part I focuses on the proposals of numerical schemes for simulations in porous and fractured media; here the challenge is to handle effectively complex geometries coupled with complex physical problems. Part II deals with some of the most advanced recent techniques, based on hybrid and extended discretization methods, for fracture and interface problems. Evolving fractures in polycrystalline materials, crack lip contact modelling, and procedures for the computation of stress intensity factors are some of the addressed topics. Part III is devoted to contributions on polygonal and polyhedral methods: these methods consist in using general polytopes, as opposed to more standard tetrahedra and hexahedra, for the discretization of partial differential equations. Specialized forms of the discontinuous Galerkin method and the virtual element method represent the core of this section. In Part IV, recent advances involving extended finite element methods and fictitious domain methods are introduced; the goals are especially to overcome some fundamental problems relating to these methods, which are basically integration at the element level and ill-conditioning of the resulting system of equations.

A common feature of all the selected contributions is the direct link between computational methodologies and their application to different engineering topics. This mix of theory and application reveals an underlying cooperation between mathematicians and engineers and highlights the way in which the scientific world is reacting to the increasing demand for simulations to contribute to the development of sustainable future technologies in engineering, biomedicine, and environmental sciences.

The Editors of this volume wish to express their sincere gratitude to all the actors involved in the X-DMS 2015 Conference, from the Institutions and Sponsors to the Scientific Committee, the organizers of the minisymposia, the plenary speakers, and all the participants.

Finally, special acknowledgments are due to Prof. Luca Formaggia, Editor-in-Chief of the SEMA SIMAI Springer Series, and to the Springer Milan editorial office for offering the opportunity to compile this volume.

Torino, Italy
Ferrara, Italy
May 2016

Giulio Ventura
Elena Benvenuti

Contents

Part I Enriched Methods for Flow and Mechanics in Heterogeneous Media

Part II Enhanced Finite Element Formulations for Fracture and Interface Problems

Part III Polygonal and Polyhedral Methods

Part IV Advances in XFEM/Fictitious Domain Methods

Part I
Enriched Methods for Flow and Mechanics in Heterogeneous Media

A Mixed Finite Element Method for Modeling the Fluid Exchange Between Microcirculation and Tissue Interstitium

Domenico Notaro, Laura Cattaneo, Luca Formaggia, Anna Scotti, and Paolo Zunino

Abstract Thanks to dimensional (or topological) model reduction techniques, small inclusions in a three-dimensional (3D) continuum can be described as one-dimensional (1D) concentrated sources, in order to reduce the computational cost of simulations. However, concentrated sources lead to singular solutions that still require computationally expensive graded meshes to guarantee accurate approximation. The main computational barrier consists in the ill-posedness of restriction operators (such as the trace operator) applied on manifolds with co-dimension larger than one. We overcome the computational challenges of approximating PDEs on manifolds with high dimensionality gap by means of nonlocal restriction operators that combine standard traces with mean values of the solution on low dimensional manifolds. This new approach has the fundamental advantage of enabling the approximation of the problem using Galerkin projections on Hilbert spaces, which could not be otherwise applied because of regularity issues. This approach, previously applied to second order PDEs, is extended here to the mixed formulation of flow problems with applications to microcirculation. In this way we calculate, in the bulk and on the 1D manifold simultaneously, the approximation of velocity and pressure fields that guarantees good accuracy with respect to mass conservation.

1 Introduction

The ultimate objective of the project is to perform large scale simulations of microcirculation. In the context of blood flow, the application of geometrical model reduction techniques plays an essential role, see for example [10, 18]. In particular, small vessels embedded into a continuum can be described as one-dimensional (1D) concentrated sources, in order to reduce the computational cost of simulations.

D. Notaro • L. Cattaneo • L. Formaggia • A. Scotti • P. Zunino (✉)
MOX, Department of Mathematics, Politecnico di Milano, piazza Leonardo da Vinci 32,
20133 Milano, Italy
e-mail: domenico.notaro@mail.polimi.it; laura1.cattaneo@polimi.it; luca.formaggia@polimi.it;
anna.scotti@polimi.it; paolo.zunino@polimi.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_1

Although the coupling of three-dimensional (3D) continua with embedded (1D) networks arises in applications of paramount importance such as microcirculation, flow through perforated media and the study of reinforced materials, it has not been well investigated yet.

Two remarkable examples of methods that were previously proposed to overcome the challenges of simulating small objects into a continuum are the *immersed boundary methods* [15, 17, 22] and the *fictitious domain methods* [11, 12, 21]. Although they share some similarities with the approach that we pursue here, they have never been applied for solving coupled partial differential equations on embedded domains.

In the particular case of microcirculation, many ad-hoc approaches have been proposed. Since capillaries can be modelled as long and narrow cylindrical vessels, asymptotic expansions that exploit the large aspect ratio of the channel can be derived to approximate the fluid exchange from one capillary to the surrounding tissue. This idea has been successfully exploited to study the microvascular flow in simple arrays of capillaries [1, 8, 9]. However, vascular networks are characterized by a complex, possibly irregular geometry. The previous semi-analytic methods may be hardly applied to realistic configurations. We believe that numerical methods may override this obstacle. For example, the method of Green's functions, has been extensively applied to the study complex vascular networks of tumors [13, 19, 20].

In this work we aim to move away from ad-hoc approaches and cast the microcirculation problem into a new unified framework to formulate and approximate coupled partial differential equations (PDEs) on manifolds with heterogeneous dimensionality. The main computational barrier consists in the ill-posedness of restriction operators (such as the trace operator) applied on manifolds with co-dimension larger than one. Following the approach introduced in [6, 7, 14], we will overcome the computational challenges of approximating PDEs on manifolds with high dimensionality gap. The main idea consists of introducing nonlocal restriction operators that combine standard traces with mean values of the solution on low dimensional manifolds, in order to couple the problem solution in 3D with the one in 1D. This new approach has the fundamental advantage to enable the approximation of the problem using Galerkin projections on Hilbert spaces, which could not be otherwise applied, because of regularity issues.

Within this general framework, the specific objective of this work is to formulate the microcirculation problem as a system of coupled 1D and 3D partial differential equations governing the flow through the capillary network and the interstitial volume, respectively. In order to obtain a good approximation of pressure and velocity fields, and in particular to satisfy mass conservation, we formulate the problem in mixed form. Then, we derive a discretization method based on mixed finite elements. Before moving forward to address applications of the method to study pathologies related to microcirculation, such as cancer [3, 4, 16], we address here a thorough validation of the solver based on two benchmark problems.

2 Model Set Up

We study a mathematical model for fluid transport in a permeable biological tissue perfused by a capillary network. The domain where the model is defined is composed by two parts, Ω and Λ, denoting the interstitial volume and the capillary bed respectively. We assume that the capillaries can be described as cylindrical vessels and Λ denotes the centerline of the capillary network. The capillary radius, R, is for simplicity considered to be constant. We decompose the network Λ into individual branches Λ_i. Branches are parametrized by the arc length s_i; a tangent unit vector $\boldsymbol{\lambda}_i$ is also defined over each branch, defining in this way an arbitrary branch orientation. Differentiation over the branches is defined using the tangent unit vector as $\partial_{s_i} := \nabla \cdot \boldsymbol{\lambda}_i$ on Λ_i, i.e. ∂_{s_i} represents the projection of ∇ along λ_i. The blood flow along each branch is described by Poiseuille's law for conservation of momentum and mass:

$$\mathbf{u}_v^i = -\frac{R^2}{8\mu}\frac{\partial p_{v,i}}{\partial s_i}\boldsymbol{\lambda}_i, \quad -\pi R^2 \frac{\partial \mathbf{u}_v^i}{\partial s_i} = g_i \qquad \text{on} \quad \Lambda_i, \tag{1}$$

where g_i is the transmural flux leaving the vessel. As a consequence of the geometrical assumptions, the vessel velocity has fixed direction and unknown scalar component along the branches, namely $\mathbf{u}_v^i = u_v^i \boldsymbol{\lambda}_i$. We shall hence formulate the vessel problem using the scalar unknown u_v. The governing flow equations for the whole network Λ are obtained by summing (1) over the index i.

We consider the interstitial volume Ω as an isotropic porous medium, described by the Darcy's law, namely

$$\mathbf{u}_t = -\frac{1}{\mu}\,\mathbb{K}\,\nabla p_t, \tag{2}$$

where $\mathbf{u}_t$ is the average velocity vector in the tissue, $\mathbb{K} = k\mathbb{I}$ is the isotropic permeability tensor, μ is the viscosity of the fluid and p_t is the fluid pressure.

The coupled problem for microcirculation and interstitial flow reads as follows

$$\begin{cases} \dfrac{\mu}{k}\mathbf{u}_t + \nabla p_t = 0 & \text{in} \quad \Omega, \\ \nabla \cdot \mathbf{u}_t - f(p_t, p_v)\delta_\Lambda = 0 & \text{in} \quad \Omega, \\ \dfrac{8\mu}{R^2} u_v + \dfrac{\partial p_v}{\partial s} = 0 & \text{in} \quad \Lambda, \\ \dfrac{\partial u_v}{\partial s} + \dfrac{1}{\pi R^2} f(p_t, p_v) = 0 & \text{in} \quad \Lambda. \end{cases} \tag{3}$$

For brevity, we avoid to number each equation of systems. In the remainder, we will refer to single sub-equations within a system using letter numbering from top to bottom (which does not explicitly appear in the expression of the equation, this notation will be adopted throughout the entire manuscript), e.g. (3) a to d.

The constitutive law for blood leakage from the capillaries to the tissue is provided by means of Starling's law of filtration,

$$f(p_t, p_v) = 2\pi R L_p (p_v - \bar{p}_t), \tag{4}$$

with

$$\bar{p}_t(s) = \frac{1}{2\pi R} \int_0^{2\pi} p_t(s, \theta) R d\theta. \tag{5}$$

Before proceeding, we write the equations in dimensionless form. We choose length, velocity and pressure as primary variables for the analysis. The corresponding characteristic values are: (1) the average spacing between capillary vessels d, (2) the average velocity in the capillary bed U, and (3) the average pressure in the interstitial space P. The dimensionless groups affecting our equations are:

- $R' = \dfrac{R}{d}$, non-dimensional radius;
- $\kappa_t = \dfrac{k}{\mu} \dfrac{P}{Ud}$, hydraulic conductivity of the tissue;
- $Q = 2\pi R' L_p \dfrac{P}{U}$, hydraulic conductivity of the capillary walls;
- $\kappa_v = \dfrac{\pi R'^4}{8\mu} \dfrac{Pd}{U}$, hydraulic conductivity of the capillary bed,

and the corresponding dimensionless equations read as follows

$$\begin{cases} \dfrac{1}{\kappa_t} \mathbf{u}_t + \nabla p_t = 0 & \text{in} \quad \Omega, \\ \nabla \cdot \mathbf{u}_t - Q \left(p_v - \bar{p}_t \right) \delta_\Lambda = 0 & \text{in} \quad \Omega, \\ \dfrac{\pi R'^2}{\kappa_v} u_v + \dfrac{\partial p_v}{\partial s} = 0 & \text{in} \quad \Lambda, \\ \dfrac{\partial u_v}{\partial s} + \dfrac{1}{\pi R'^2} Q \left(p_v - \bar{p}_t \right) = 0 & \text{in} \quad \Lambda. \end{cases} \tag{6}$$

For simplicity of notation, we used the same symbols for the dimensionless variables, i.e velocities and pressure scaled by U and P, respectively.

Remark 1 Equations (6) b and d, counting from the top, can be combined up to obtain a more meaningful formulation of the mass conservation law, namely

$$\nabla \cdot \mathbf{u}_t + \pi R'^2 \frac{\partial u_v}{\partial s} \delta_\Lambda = 0 \qquad \text{in} \quad \Omega, \tag{7}$$

meaning that the total amount of fluid in the domain $\Omega \cup \Lambda$ must be preserved.

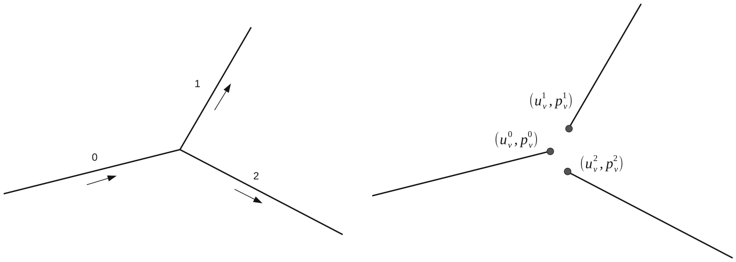


Fig. 1 On the *left*, a simple network made by a single Y-shaped bifurcation. *Arrows* show the flow orientation of one inflow branch on the left of the bifurcation point and two outflow branches on the right. On the *right*, the discretization of vessels network is shown. The domain has been split into branches, the flow problem is defined over each branch and compatibility conditions are enforced at the junction point

Boundary conditions will be specified further on for both the tissue and vessel problems. The imposition of suitable compatibility conditions at the bifurcations or branching points of the capillary tree is also necessary to guarantee well posedness of (6). Specifically, we shall enforce conservation of mass and continuity of total pressure at junctions. Let us introduce these conditions in a simple Y-shaped bifurcation network (Fig. 1). Since in the reduced 1D model of the capillary network the cross-section is supposed to be constant over the whole network, the conservation of flow rate is equivalent to require that in correspondence of the junction point $\mathbf{x}_M$ the inflow velocity u_v^0 is equal the sum of the outflow velocities u_v^1, u_v^2, namely $u_v^0(\mathbf{x}_M) = u_v^1(\mathbf{x}_M) + u_v^2(\mathbf{x}_M)$. Similarly, we require the pressure over each branch to be the same at the junction, namely $p_v^0(\mathbf{x}_M) = p_v^1(\mathbf{x}_M) = p_v^2(\mathbf{x}_M)$. The general case of an arbitrary number of *critical* points, possibly with different number of inflow and outflow branches, will be described in Sect. 3. Indeed, it is important to emphasize that such compatibility conditions will be enforced in a natural way, at the level of the variational formulation.

3 Variational Formulation

In order to obtain the weak formulation of the tissue interstitium problem, we multiply Eqs. (6) a and b, counting from the top with sufficiently smooth functions and integrate over the volume Ω, namely

$$\int_{\Omega} \frac{1}{\kappa_t} \mathbf{u}_t \cdot \mathbf{v}_t \, d\mathbf{x} \; + \; \int_{\Omega} \nabla p_t \cdot \mathbf{v}_t \, d\mathbf{x} \; = \; 0, \tag{8}$$

$$\int_{\Omega} (\nabla \cdot \mathbf{u}_t) \; q_t \, d\mathbf{x} \; - \; \int_{\Omega} Q \, (p_v - \bar{p}_t) \, \delta_{\Lambda} \, q_t \, d\mathbf{x} \; = \; 0. \tag{9}$$

We now apply the Green's theorem to (8) to obtain an anti-symmetric formulation of the Darcy's problem in the tissue:

$$\int_{\Omega} \frac{1}{\kappa_t} \mathbf{u}_t \cdot \mathbf{v}_t \, d\mathbf{x} \; - \; \int_{\Omega} p_t \, (\nabla \cdot \mathbf{v}_t) \;\, d\mathbf{x} \; + \; \int_{\partial\Omega} p_t \, \mathbf{v}_t \cdot \mathbf{n} \, d\sigma(\mathbf{x}) \; = \; 0, \tag{10}$$

$$\int_{\Omega} (\nabla \cdot \mathbf{u}_t) \; q_t \, d\mathbf{x} \; - \; \int_{\Omega} Q \, (p_v - \bar{p}_t) \, \delta_{\Lambda} \, q_t \, d\mathbf{x} \; = \; 0. \tag{11}$$

As concerns the choice of boundary conditions, for simplicity, we enforce a given pressure distribution over $\partial\Omega$, namely

$$p_t = g_t \quad \text{on} \quad \partial\Omega, \tag{12}$$

where $g_t \in L^2(\partial\Omega)$. The weak formulation of the problem in Ω reads

$$\int_{\Omega} \frac{1}{\kappa_t} \mathbf{u}_t \cdot \mathbf{v}_t \, d\mathbf{x} \; - \; \int_{\Omega} p_t \, (\nabla \cdot \mathbf{v}_t) \;\, d\mathbf{x} \; = \; - \int_{\partial\Omega} g_t \, \mathbf{v}_t \cdot \mathbf{n} \, d\sigma(\mathbf{x})$$

$$\int_{\Omega} (\nabla \cdot \mathbf{u}_t) \; q_t \, d\mathbf{x} \; - \; \int_{\Omega} Q \, (p_v - \bar{p}_t) \, \delta_{\Lambda} \, q_t \, d\mathbf{x} \; = \; 0 \quad .$$

For the vessel problem we start giving a general functional framework. At this point, we only require regularity for vessel velocity and pressure over each branch separately:

$$V_v = \bigcup_{i=1}^{N} H^1(\Lambda_i) \qquad Q_v = \bigcup_{i=1}^{N} L^2(\Lambda_i).$$

The definition of trial and test spaces will be revised in the sequel, in the light of the particular junction conditions we will chose, while no boundary conditions are enforced in the definition of the spaces. As for the tissue problem, we multiply Eqs. (6.c),(6.d) by sufficiently smooth test functions and integrate over Λ:

$$\int_{\Lambda} \frac{\pi R'^2}{\kappa_v} u_v \, v_v \, ds \; + \; \int_{\Lambda} \frac{\partial p_v}{\partial s} \, v_v \, ds = 0, \tag{13}$$

$$\int_{\Lambda} \frac{\partial u_v}{\partial s} \, q_v \, ds \; + \; \frac{1}{\pi R'^2} \int_{\Lambda} Q \, (p_v - \bar{p}_t) \; q_v \, ds = 0. \tag{14}$$

The integration by parts is not trivial in this case because the vessel variables p_v and u_v *may* be discontinuous at multiple junctions. Let us treat separately the second integral of (13) and decompose it over the individual branches Λ_i:

$$\int_{\Lambda} \frac{\partial p_v}{\partial s} \, v_v \, ds = \sum_{i=1}^{N} \int_{\Lambda_i} \frac{\partial p_v}{\partial s} \, v_v \, ds = - \int_{\Lambda} p_v \, \frac{\partial v_v}{\partial s} \, ds \; + \; \sum_{i=1}^{N} [p_v \, v_v \,]_{\Lambda_i^-}^{\Lambda_i^+} \; , \tag{15}$$

where Λ_i^- and Λ_i^+ represent the inflow and outflow boundaries of Λ_i, according to the orientation $\boldsymbol{\lambda}_i$. Let us define the set of the indexes of junction points:

$$\mathscr{J} := \left\{ j \in \mathbb{N} \ : \ s_j \in \Lambda, \ \#(P_{s_j}) \geq 2 \right\},$$

where P_{s_j} is the *patch* of the j-th junction node, i.e. the collection of all branches joining at the node, and # indicates the counting measure. We also need the following disjoint partition of the indexes in P_{s_j}. According to the orientation unit vector $\boldsymbol{\lambda}_i$, for any branching point s_j we distinguish branches that are entering the node, whose contribution to mass conservation is positive, from branches who are leaving the node, whose contribution is negative. The former are branches whose outflow region coincides with the point s_j, while for the latter it is the inflow region:

$$\mathscr{P}_j^{out} := \left\{ i \in \{1, \ldots, N\} \ : \ \Lambda_i^+ \equiv \{s_j\} \right\},$$

$$\mathscr{P}_j^{in} := \left\{ i \in \{1, \ldots, N\} \ : \ \Lambda_i^- \equiv \{s_j\} \right\},$$

for all $j \in \mathscr{J}$. At this point, the fluid mass conservation at each node can be expressed as follows

$$\sum_{i \in \mathscr{P}_j^{out}} u_v \,|_{\Lambda_i^+} - \sum_{i \in \mathscr{P}_j^{in}} u_v \,|_{\Lambda_i^-} = 0, \quad \forall j \in \mathscr{J}, \tag{16}$$

where $\Lambda^{in}, \Lambda^{out}$ indicate the collection of inflow and outflow boundaries of the vessel network, i.e. non junction points where the tangent unit vector is inward-pointing and outward-pointing, respectively. This collection contains the boundary points, i.e. the extrema that also belong to $\partial\Omega$, but the inclusion may be strict. However, in this contribution we do not address the issue of network extrema belonging to $\overset{\circ}{\Omega}$, i.e. we do not consider immersed tips.

In order to enforce such conditions, we proceed as follows. First, we reformulate the last term in (15) by isolating the terms relative to inflow junction nodes from those relative to outflow nodes, namely

$$\sum_{i=1}^{N} [p_v \, v_v]_{\Lambda_i^-}^{\Lambda_i^+} = \sum_{j \in \mathscr{J}} \left[\sum_{i \in \mathscr{P}_j^{out}} p_v \, v_v \,|_{\Lambda_i^+} - \sum_{i \in \mathscr{P}_j^{in}} p_v \, v_v \,|_{\Lambda_i^-} \right] + [p_v \, v_v]_{\Lambda^{in}}^{\Lambda^{out}} .$$

Here, we have implicitly assumed the trace of $(p_v \, v_v)$ over Λ_i exists for v_v smooth enough, i.e. the evaluation of the product at the extrema of Λ_i makes sense. Furthermore, we write $(p_v \, v_v)\,(s_j) = p_v(s_j)\; v_v(s_j)$ for some point $s_j \in \Lambda$. This is feasible if the trace of the pressure exists. Obviously, a general L^2 function is not sufficient, the natural choice is $p_v \in \mathscr{C}^0(\bar{\Lambda})$, that in particular implies compatibility of pressure values at the junctions. Indeed, if the pressure is continuous at the junction, we have

$$p_v\,|_{\Lambda_i^-} \equiv p_v(s_j) \equiv p_v\,|_{\Lambda_k^+} \quad \forall i \in \mathscr{P}_j^{in},\; k \in \mathscr{P}_j^{out} \quad \forall j \in \mathscr{J}. \tag{17}$$

Under that hypothesis, we finally factorize out the pressure and isolate a term that corresponds to the junction conditions for the velocity test functions, that is

$$\sum_{j\in\mathscr{J}} p_v(s_j)\left[\sum_{i\in\mathscr{P}_j^{out}} v_v\,|_{\Lambda_i^+} - \sum_{i\in\mathscr{P}_j^{in}} v_v\,|_{\Lambda_i^-}\right].$$

Then, we weakly enforce mass conservation into the variational formulation by multiplying (16) by the pressure test functions q_v, which act as a Lagrange multiplier for this constraint, namely

$$\sum_{j\in\mathscr{J}} q_v(s_j)\left[\sum_{i\in\mathscr{P}_j^{out}} u_v\,|_{\Lambda_i^+} - \sum_{i\in\mathscr{P}_j^{in}} u_v\,|_{\Lambda_i^-}\right].$$

Finally, after adding the previous term to Eq. (14), the weak formulation of the vessel problem reads

$$\begin{aligned}\int_\Lambda \frac{\pi R'^2}{\kappa_v} u_v\, v_v\, ds \;-\; \int_\Lambda p_v \frac{\partial v_v}{\partial s}\, ds \;+\; [p_v\, v_v]_{\Lambda^{in}}^{\Lambda^{out}}&\\ +\; \sum_{j\in\mathscr{J}} p_v(s_j)\left[\sum_{i\in\mathscr{P}_j^{out}} v_v\,|_{\Lambda_i^+} - \sum_{i\in\mathscr{P}_j^{in}} v_v\,|_{\Lambda_i^-}\right] \;&=\; 0,\end{aligned} \tag{18}$$

$$\begin{aligned}\int_\Lambda \frac{\partial u_v}{\partial s} q_v\, ds \;+\; \frac{1}{\pi R'^2}\int_\Lambda Q\,(p_v - \bar{p}_t)\; q_v\, ds&\\ -\; \sum_{j\in\mathscr{J}} q_v(s_j)\left[\sum_{i\in\mathscr{P}_j^{out}} u_v\,|_{\Lambda_i^+} - \sum_{i\in\mathscr{P}_j^{in}} u_v\,|_{\Lambda_i^-}\right] \;&=\; 0.\end{aligned} \tag{19}$$

Concerning the boundary conditions for the vessels network, the natural choice is to enforce a given pressure distributions at the inflow and the outflow of the network, $p_v = g_v$ on $\Lambda^{in} \cup \Lambda^{out}$. The generic regularity requirements for the Dirichlet's datum are measurability and square-summability, namely $g_v \in L^2(\Lambda^{in} \cup \Lambda^{out})$. In practice, we consider a constant pressure drop $\Delta P_v = P_v^{out} - P_v^{in}$:

$$g_v(s) = \begin{cases} P_v^{in} & s \in \Lambda^{in} \\ P_v^{out} & s \in \Lambda^{out}. \end{cases} \tag{20}$$

Since we are considering the mixed formulation of the problem, we enforce such condition in a weak *natural* way.

At this point, we combine (10), (11), (18), (19) to obtain the whole weak formulation of our 3D-1D coupled model of fluid exchange between microcirculation and tissue interstitium. The variational formulation of problem (6) consists of finding $\mathbf{u}_t \in \mathbf{V}_t\,,\ p_t \in Q_t\,,\ u_v \in V_v\,,\ p_v \in Q_v$ s.t.

$$\begin{cases} \dfrac{1}{\kappa_t}\big(\mathbf{u}_t\,,\,\mathbf{v}_t\big)_\Omega \;-\; \big(p_t\,,\,\nabla\cdot\mathbf{v}_t\big)_\Omega \;=\; -\,\big(g_t\,,\,\mathbf{v}_t\cdot\mathbf{n}\big)_{\partial\Omega} & \forall \mathbf{v}_t \in \mathbf{V}_t, \\ \big(\nabla\cdot\mathbf{u}_t\,,\,q_t\big)_\Omega \;-\; Q\big(\,(p_v - \bar{p}_t)\,\delta_\Lambda\,,\,q_t\big)_\Omega \;=\; 0 & \forall q_t \in Q_t, \\ \dfrac{\pi R'^2}{\kappa_v}\big(u_v\,,\,v_v\big)_\Lambda \;-\; \big(p_v\,,\,\partial_s v_v\big)_\Lambda & \\ \qquad +\; \sum_j p_v(s_j)\Big[\sum_i v_v\,|_{\Lambda_i^+} - \sum_i v_v\,|_{\Lambda_i^-}\Big] = \;-\,[\,g_v\,v_v\,]_{\Lambda^{in}}^{\Lambda^{out}} & \forall v_v \in V_v, \\ \big(\partial_s u_v\,,\,q_v\big)_\Lambda \;+\; \dfrac{1}{\pi R'^2}\,Q\big(p_v - \bar{p}_t\,,\,q_v\big)_\Lambda & \\ \qquad -\; \sum_j q_v(s_j)\Big[\sum_i u_v\,|_{\Lambda_i^+} \;-\; \sum_i u_v\,|_{\Lambda_i^-}\Big] \;=\; 0 & \forall q_v \in Q_v. \end{cases} \tag{21}$$

4 Numerical Approximation

The discretization of problem (6) is achieved by means of the finite element method that arises from the variational formulation (21) combined with a discretization of the domain. In particular, one of the advantage of our formulation is that the partitions of Ω and Λ are completely independent. Let us now analyze the two approximations separately.

We denote with $\mathscr{T}_t^h$ an admissible family of partitions of $\bar{\Omega}$ into tetrahedrons K

$$\bar{\Omega} = \bigcup_{K \in \mathscr{T}_t^h} K,$$

that satisfies the usual conditions of a conforming triangulation of Ω. Here, h denotes the mesh characteristic size, i.e. $h = \max_{K \in \mathscr{T}_t^h} k_K$, being h_K the diameter of simplex K. Moreover, we are implicitly assuming that Ω is a *polygonal* domain. The solutions of (21) a and b, counting from the top are approximated using discontinuous piecewise-polynomial finite elements for pressure and

$\mathbf{H}^{div}$-conforming *Raviart-Thomas* finite elements [2] for velocity, namely

$$\mathbb{Y}_h^k := \left\{ w_h \in L^2(\Omega) \ : \ w_h|_K \in \mathscr{P}_{k-1}(K) \quad \forall K \in \mathscr{T}_t^h \right\},$$

$$\mathbb{RT}_h^k := \left\{ \mathbf{w}_h \in \mathbf{H}\left((div, \Omega\right) \ : \ \mathbf{w}_h|_K \in \mathscr{P}_{k-1}(K; \mathbb{R}^d) \oplus \mathbf{x}\, \mathscr{P}_{k-1}(K) \quad \forall K \in \mathscr{T}_t^h \right\},$$

for every integer $k \geq 0$, where $\mathscr{P}_k$ indicates the standard space of polynomials of degree $\leq k$ in the variables $\mathbf{x} = (x_1, \ldots, x_d)$. For the simulations presented later on, the lowest order *Raviart-Thomas* approximation has been adopted, corresponding to $k = 1$ above. In numerical experiments performed on the 3D problem alone (the test case is not reported here), we have observed quadratic convergence of the pressure field and linear convergence of the velocity field.

Concerning the capillary network, we adopt the same domain splitting technique described at the continuous level, obtaining the following discrete domain:

$$\Lambda_h = \bigcup_{i=1}^{N} \Lambda_i^h,$$

where Λ_i^h is a finite element mesh on the one-dimensional manifold Λ_i, i.e. a partition of the i-th network branch made by a sufficiently large number of segments.

The solution of sub-equations (21) c and (21) d, counting from the top, over a given branch Λ_i is approximated using continuous piecewise-polynomial finite element spaces for both pressure and velocity. Since we want the vessel velocity to be discontinuous at multiple junctions, we define the related finite element space over the whole network as the collection of the local spaces of the single branches. Conversely, the pressure has been assumed to be continuous over the network. We will use the following families of finite element spaces for pressure and velocity, respectively:

$$\mathbb{X}_h^{k+1}(\Lambda) := \left\{ w_h \in \mathscr{C}^0(\bar{\Lambda}) \ : \ w_h|_S \in \mathscr{P}_k(S) \quad \forall S \in \Lambda^h \right\},$$

$$\mathbb{W}_h^{k+2}(\Lambda) := \bigcup_{i=1}^{N} \mathbb{X}_h^{k+1}(\Lambda_i),$$

for every integer $k \geq 0$. As a result, we use generalized Taylor-Hood elements on each network branch, satisfying in this way the local stability of the mixed finite element pair for the network. At the same time, we guarantee that the pressure approximation is continuous over the entire network Λ. In particular, for the numerical experiments shown later on we have used the lowest order, that is $k = 1$.

The discrete formulation arising from (21) is hence easily obtained by adding the subscript h to the weak continuous formulation: find $\mathbf{u}_{t,h} \in \mathbf{V}_t^h$, $p_{t,h} \in Q_t^h$, $u_{v,h} \in V_v^h$, $p_{v,h} \in Q_v^h$ s.t.

$$\begin{cases} \dfrac{1}{\kappa_t}\big(\mathbf{u}_{t,h}\,,\,\mathbf{v}_{t,h}\big)_\Omega \;-\; \big(p_{t,h}\,,\,\nabla\cdot\mathbf{v}_{t,h}\big)_\Omega \;=\; -\;\big(g_{t,h}\,,\,\mathbf{v}_{t,h}\cdot\mathbf{n}\big)_{\partial\Omega} & \forall\mathbf{v}_{t,h}\in\mathbf{V}_t^h, \\ \big(\nabla\cdot\mathbf{u}_{t,h}\,,\,q_{t,h}\big)_\Omega \;-\; Q\big(\,(p_{v,h}-\bar{p}_{t,h})\,\delta_\Lambda\,,\,q_{t,h}\big)_\Omega \;=\; 0 & \forall q_{t,h}\in Q_t^h, \\ \dfrac{\pi R'^2}{\kappa_v}\big(u_{v,h}\,,\,v_{v,h}\big)_\Lambda \;-\; \big(p_{v,h}\,,\,\partial_s v_{v,h}\big)_\Lambda & \\ \quad+\;\sum_j p_{v,h}(s_j)\Big[\sum_i v_{v,h}\,|_{\Lambda_i^+} - \sum_i v_{v,h}\,|_{\Lambda_i^-}\Big] \;=\; -\,[\,g_{v,h}\,v_{v,h}\,]_{\Lambda^{in}}^{\Lambda^{out}} & \forall v_{v,h}\in V_v^h \\ \big(\partial_s u_{v,h}\,,\,q_{v,h}\big)_\Lambda \;+\; \dfrac{1}{\pi R'^2}\,Q\big(p_{v,h}-\bar{p}_{t,h}\,,\,q_{v,h}\big)_\Lambda & \\ \quad-\;\sum_j q_{v,h}(s_j)\Big[\sum_i u_{v,h}\,|_{\Lambda_i^+} \;-\; \sum_i u_{v,h}\,|_{\Lambda_i^-}\Big] = \;0 & \forall q_{v,h}\in Q_v^h, \end{cases} \tag{22}$$

where $g_{t,h}$, $g_{v,h}$ indicate the discrete counterparts of continuous boundary data.

We observe that (22) is a generalized saddle-point problem arising from the combination of local problems with mass conservation constraints (see also (23)), such as the mixed formulation of Darcy equation and the incompressible flow on each network branch with junction conditions. Although, we guarantee local stability of each block, the global well-posedness is an still an open problem, which is under investigation.

4.1 Algebraic Formulation

Let us now derive the algebraic form of our discrete problem. We define the number of degrees of freedom of our discrete (finite) spaces as:

$$\begin{aligned} N_t^h &:= dim\left(\mathbf{V}_t^h\right), \qquad & M_t^h &:= dim\left(Q_t^h\right), \\ N_v^h &:= dim\left(V_v^h\right), \qquad & M_v^h &:= dim\left(Q_v^h\right) \quad . \end{aligned}$$

We denote with $\{\boldsymbol{\varphi}_t^i\}_{i=1}^{N_t^h} \times \{\psi_t^i\}_{i=1}^{M_t^h}$ and $\{\varphi_v^i\}_{i=1}^{N_v^h} \times \{\psi_v^i\}_{i=1}^{M_v^h}$ the finite element basis for $\mathbf{V}_t^h \times Q_t^h$ and $V_v^h \times Q_v^h$ respectively. These two sets are completely independent, since the 3D and 1D meshes do not conform. We set:

$$\mathbf{u}_t^h(\mathbf{x}) = \sum_{j=1}^{N_t^h} U_t^j \boldsymbol{\varphi}_t^j(\mathbf{x}), \quad p_t^h(\mathbf{x}) = \sum_{j=1}^{M_t^h} P_t^j \psi_t^j(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega_t,$$

$$u_v^h(s) = \sum_{j=1}^{N_v^h} U_v^j \varphi_v^j(s), \quad p_v^h(s) = \sum_{j=1}^{M_v^h} P_v^j \psi_v^j(s) \qquad \forall s \in \Lambda,$$

being $\mathbf{U}_t = \{U_t^j\}_{j=1}^{N_t^h}$, $\mathbf{P}_t = \{P_t^j\}_{j=1}^{M_t^h}$, $\mathbf{U}_v = \{U_v^j\}_{j=1}^{N_v^h}$ and $\mathbf{P}_v = \{P_v^j\}_{j=1}^{M_v^h}$, the degrees of freedom of the finite element approximation. Then, by replacing the linear combinations within the discrete weak form (22) and using the linearity of the inner product, from (22) we deduce the following linear system:

$$\begin{bmatrix} \mathbb{M}_{tt} & -\mathbb{D}_{tt}^T & \mathbb{O} & \mathbb{O} \\ \mathbb{D}_{tt} & \mathbb{B}_{tt} & \mathbb{O} & -\mathbb{B}_{tv} \\ \mathbb{O} & \mathbb{O} & \mathbb{M}_{vv} & -\mathbb{D}_{vv}^T - \mathbb{J}_{vv}^T \\ \mathbb{O} & -\mathbb{B}_{vt} & \mathbb{D}_{vv} + \mathbb{J}_{vv} & \mathbb{B}_{vv} \end{bmatrix} \begin{bmatrix} \mathbf{U}_t \\ \mathbf{P}_t \\ \mathbf{U}_v \\ \mathbf{P}_v \end{bmatrix} = \begin{bmatrix} \mathbf{F}_t \\ \mathbf{0} \\ \mathbf{F}_v \\ \mathbf{0} \end{bmatrix}. \tag{23}$$

Standard finite element matrices and right hand sides are defined as follows

$$[\mathbb{M}_{tt}]_{i,j} := \frac{1}{\kappa_t}\left(\boldsymbol{\varphi}_t^j, \boldsymbol{\varphi}_t^i\right)_\Omega \qquad \mathbb{M}_{tt} \in \mathbb{R}^{N_t^h \times N_t^h},$$

$$[\mathbb{D}_{tt}]_{i,j} := \left(\nabla \cdot \boldsymbol{\varphi}_t^j, \psi_t^i\right)_\Omega \qquad \mathbb{D}_{tt} \in \mathbb{R}^{N_t^h \times M_t^h},$$

$$[\mathbb{D}_{vv}]_{i,j} := \left(\partial_s \varphi_v^j, \psi_v^i\right)_\Lambda \qquad \mathbb{D}_{vv} \in \mathbb{R}^{N_v^h \times M_v^h},$$

$$[\mathbb{M}_{vv}]_{i,j} := \pi R'^2/\kappa_v \left(\varphi_v^j, \varphi_v^i\right)_\Lambda \qquad \mathbb{M}_{vv} \in \mathbb{R}^{N_v^h \times N_v^h},$$

$$[\mathbf{F}_t]_i := -\left(g_{t,h}, \boldsymbol{\varphi}_t^i \cdot \boldsymbol{n}\right)_{\partial\Omega} \qquad \mathbf{F}_t \in \mathbb{R}^{N_t^h},$$

$$[\mathbf{F}_v]_i := -[g_{v,h}\, \varphi_v^i]_{\Lambda^{in}}^{\Lambda^{out}} \qquad \mathbf{F}_v \in \mathbb{R}^{N_v^h}.$$

For the implementation of exchange matrices, namely $\mathbb{B}_{tt}, \mathbb{B}_{tv}, \mathbb{B}_{vt}, \mathbb{B}_{vv}$, we define two discrete operators: the first one extracts the mean value of a generic basis function of Q_t^h, while the second interpolates between Q_t^h and Q_v^h. For every node $s_k \in \Lambda^h$ we define $\mathcal{T}_\gamma(s_k)$ as the discretization of the perimeter of the vessel $\gamma(s_k)$, see Fig. 2 for an illustration. For simplicity, we assume that $\gamma(s_k)$ is a circle of radius R defined on the orthogonal plane to Λ^h at point s_k. The set of points of $\mathcal{T}_\gamma(s_k)$ is used to interpolate the basis functions ψ_t^i. Let us introduce a local discrete interpolation matrix $\boldsymbol{\Pi}_\gamma(s_k)$ which returns the values of each test function ψ_t^i on the set of points belonging to $\mathcal{T}_\gamma(s_k)$. Then, we consider the average operator

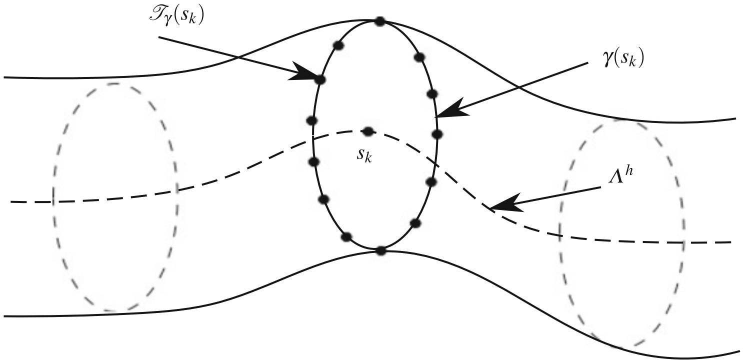


Fig. 2 Illustration of the vessel with its centerline Λ^h, a cross section, its perimeter $\gamma(s_k)$ and its discretization $\mathscr{T}_\gamma(s_k)$ used for the definition of the interface operators $\bar{\pi}_{vt} : Q_t^h \longrightarrow Q_v^h$ and $\pi_{tv} : Q_v^h \longrightarrow Q_t^h$

$\bar{\pi}_{vt} : Q_t^h \longrightarrow Q_v^h$ such that $\bar{q}_t = \bar{\pi}_{vt} q_t$. The matrix $\bar{\boldsymbol{\Pi}}_{vt}$ that corresponds to this operator belongs to $\mathbb{R}^{M_v^h \times M_t^h}$ and it is constructed such that each row is defined as,

$$\bar{\boldsymbol{\Pi}}_{vt}|_k = \mathbf{w}^T(s_k)\, \boldsymbol{\Pi}_\gamma(s_k) \qquad k = 1, \ldots, M_v^h \tag{24}$$

where $\mathbf{w}$ are the weights of the quadrature formula used to approximate the integral

$$\bar{q}_t(s) = \frac{1}{2\pi R} \int_0^{2\pi} q_t(s, \theta)\, R\, d\theta$$

on the nodes belonging to $\mathscr{T}_\gamma(s_k)$. The discrete interpolation operator $\pi_{tv} : Q_v^h \longrightarrow Q_t^h$ returns the value of each basis function belonging to Q_t^h in correspondence of nodes of Q_v^h. In algebraic form it is expressed as an interpolation matrix $\boldsymbol{\Pi}_{tv} \in \mathbb{R}^{M_v^h \times M_t^h}$. Using these tools we obtain:

$$\mathbb{B}_{tt} = Q\, \boldsymbol{\Pi}_{vt}^T\, \mathbb{M}_{vv}^P\, \bar{\boldsymbol{\Pi}}_{vt}, \tag{25}$$

$$\mathbb{B}_{tv} = Q\, \boldsymbol{\Pi}_{vt}^T\, \mathbb{M}_{vv}^P, \tag{26}$$

$$\mathbb{B}_{vt} = Q/\pi R'^2\, \mathbb{M}_{vv}^P\, \bar{\boldsymbol{\Pi}}_{vt}, \tag{27}$$

$$\mathbb{B}_{vv} = Q/\pi R'^2\, \mathbb{M}_{vv}^P, \tag{28}$$

being $\mathbb{M}_{vv}^P$ the pressure mass matrix for the vessel problem defined by

$$\left[\mathbb{M}_{vv}^P\right]_{i,j} := \left(\psi_v^j\,,\, \psi_v^i\right)_\Lambda.$$

Concerning the implementation of junction compatibility conditions, we introduce a linear operator giving the restriction with sign of a basis function of V_v^h over a given junction node. For a given $k \in \mathscr{J}$, we define $\mathscr{R}_k : V_v^h \longrightarrow \mathbb{R}$ such that:

$$\mathscr{R}_k(\varphi_v^j) := \begin{cases} +\varphi_v^j(s_k) \; j \; in \; \Lambda_l^h \; \wedge \; l \in \mathscr{P}_k^{out} \\ -\varphi_v^j(s_k) \; j \; in \; \Lambda_l^h \; \wedge \; l \in \mathscr{P}_k^{in} \end{cases} \tag{29}$$

for all $j = 1, \ldots, N_v^h$, where the expression "$j \; in \; \Lambda_l^h$" means that the j-th dof is linked to some vertex of the l-th branch. Note that we are implicitly using the usual property of Lagrangian finite element basis functions, i.e. that they vanish on all nodes except the related one. As a consequence, our definition is consistent for all junction vertexes. Indeed, $\mathscr{R}_k$ may only assume values $-1, 0, +1$ and in particular $\mathscr{R}_k(\varphi_v^j) = 0$ for all couples of indexes (k, j) that are uncorrelated. Furthermore, the definition of $\mathscr{R}_k$ can be trivially extended to all network vertexes. Using this operator, the generic (i, j) element of $\mathbb{J}_{vv}$ may be computed as follows

$$[\mathbb{J}_{vv}]_{i,j} = -\sum_{k \in \mathscr{J}} \mathscr{R}_k(\varphi_v^j) \, \psi_v^i(s_k) \,. \tag{30}$$

5 Numerical Experiments

We validate the mixed-finite element solver through the following test cases, illustrated in Fig. 3, which have been designed to obtain sufficient generality with a straightforward interpretation of the results:

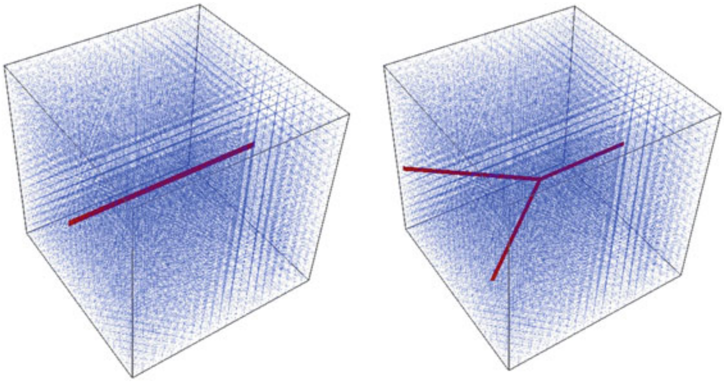


Fig. 3 (*Left*) Computational domain for test-case I. The discrete network Λ_h is made by a single capillary vessel immersed in a unitary slab of tissue interstitium, Ω_h. We have used a discretization step $h = 0.05$ for both the 1D and 3D problems. (*Right*) Computational domain for test-case II. The discrete vessels network Λ_h is made by three capillaries joined junction point $\mathbf{x}_M = (0.5, 0.5, 0.5)$: Λ_h^0 entering branch, Λ_h^1 and Λ_h^2 exiting branches. The tissue interstitium domain Ω_h is a unit cube. Again, we have used a discretization step $h = 0.05$ for both the 1D and 3D problems

(1) Coupled 3D-1D problem on a single branch;
(2) Coupled 3D-1D problem on a Y-shaped bifurcation.

In this way, we address the two main modeling issues: (1) first, we test the ability of the computational model to approximate the coupling between 3D and 1D equations; (2) second, we verify that the assembly of junction conditions works properly.

5.1 Coupled 3D-1D Problem on a Single Branch

For such a simple setting, we can easily isolate the exchange terms. The 3D-1D coupled problem is given by (6). In this case, the integration by parts in (6)(iii) is standard since there are not any junction points. As a consequence, we replace condition (15) with the following:

$$\begin{aligned}\int_{\Lambda} \frac{\partial p_v}{\partial s}\, v_v \, ds &= -\int_{\Lambda} p_v \, \frac{\partial v_v}{\partial s}\, ds \;+\; [p_v \, v_v\,]_{\Lambda^{in}}^{\Lambda^{out}} \\ &= -\int_0^1 p_v \, \frac{\partial v_v}{\partial s}\, ds \;+\; p_v(1)\, v_v(1) - p_v(0)\, v_v(0) \quad .\end{aligned}$$

Therefore, we obtain the following linear system:

$$\begin{bmatrix} \mathbb{M}_{tt} & -\mathbb{D}_{tt}^T & \mathbb{O} & \mathbb{O} \\ \mathbb{D}_{tt} & \mathbb{B}_{tt} & \mathbb{O} & -\mathbb{B}_{tv} \\ \mathbb{O} & \mathbb{O} & \mathbb{M}_{vv} & -\mathbb{D}_{vv}^T \\ \mathbb{O} & -\,\mathbb{B}_{vt} & \mathbb{D}_{vv} & \mathbb{B}_{vv} \end{bmatrix} \begin{bmatrix} \mathbf{U}_t \\ \mathbf{P}_t \\ \mathbf{U}_v \\ \mathbf{P}_v \end{bmatrix} = \begin{bmatrix} \mathbf{F}_t \\ \mathbf{0} \\ \mathbf{F}_v \\ \mathbf{0} \end{bmatrix} . \tag{31}$$

We recall that submatrices in (31) have been defined in Sect. 4.1. Nevertheless, according to the above expression of vessel boundary term it is possible to specify the right hand side, namely

$$\mathbf{F}_v \;=\; -[\, g_{v,h}\, \varphi_v^i\,]_0^1 \;\equiv\; \begin{bmatrix} g_{v,h}(0) \\ 0 \\ \vdots \\ 0 \\ -g_{v,h}(1) \end{bmatrix} \tag{32}$$

being $g_{v,h}$ the discrete counterpart of the vessel boundary datum. In the last equality we used the fundamental property of finite element basis functions. Note that (31) equals the generic linear system (23) in the special case $\mathbb{J}_{vv} = \mathbb{O}$.

5.1.1 Numerical Results

For the tissue sample Ω_h we use a tetrahedral structured mesh, $\mathscr{T}_h$, with characteristic size $h = 1/20$; the same step has been used for the network discretization Λ_h, resulting in 48,000 elements for the approximation of interstitial volume and 60 elements for the discrete network. We prescribe the following boundary conditions:

$$p_t|_{\partial\Omega} = 0\,, \quad p_v(0) = 1.0\,, \quad p_v(1) = 0.5\,. \tag{33}$$

For the solution of the linear system (31) we developed a C++ code based on GetFEM++ (see https://home.gna.org/getfem), an open-source general purpose finite element library. Specifically, we applied the direct solver *SuperLU* 3.0 (see http://crd.lbl.gov/verb~xiaoye/SuperLU). Numerical solutions are shown in Fig. 4. These plots show qualitatively that the definition and implementation of the method works properly. In order to find a *quantitative* way to validate our numerical method we exploit the exact solution proposed by Chapman and Shipley [5] for the single

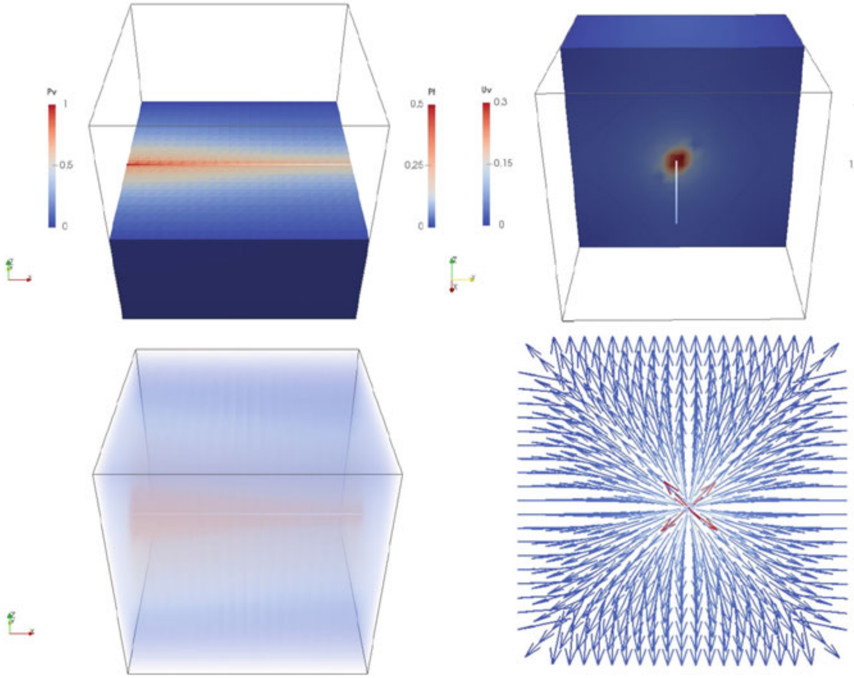


Fig. 4 Coupling between the vessel and tissue interstitium. Numerical solutions obtained with mesh size $h = 0.05$ and parameters $\kappa_t = \kappa_v = 1, R' = 1, Q = 1$. On the left a double-check for pressure exchange: (*top-left*) visualization of tissue and vessel pressures at the transversal medium plane, (*bottom-left*) a 3D qualitative representation. On the right the velocity exchange: (*top-right*) visualization of tissue and vessel velocities at the axial medium plane, (*bottom-right*) a 2D view of the vector field

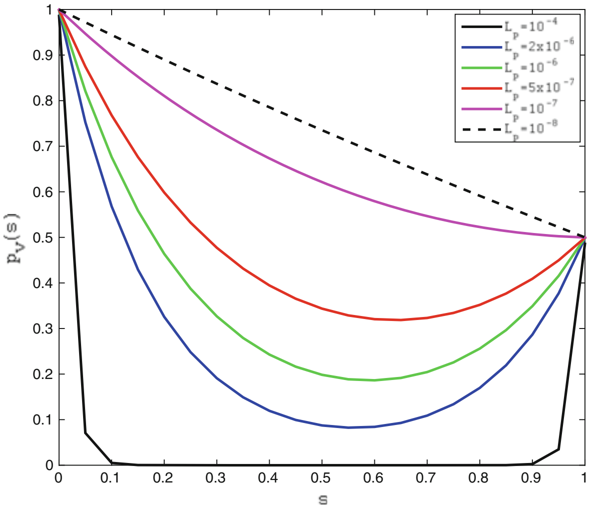


Fig. 5 Capillary pressure as function of arclength s for different vascular permeabilities $\hat{L}_p = 10^{-4}, 2 \times 10^{-6}, 10^{-6}, 5 \times 10^{-7}, 10^{-7}, 10^{-8}$. To be compared with Fig. 7 in [5]

branch problem. In that work the authors model a fluid flow through the leaky neovasculature and porous interstitium of a solid tumor, in particular they consider the simplest case of an isolated capillary immersed in a tumor tissue, giving rise to the same problem addressed here.

Finally, in order to reproduce numerical results of [5] we choose the non-dimensional parameters of the problem as follows

$$R' = 10^{-2}\ , \quad \kappa_t = 4\ , \quad \kappa_v = \pi R'^3/8\hat{L}_p\ , \quad Q = 2\pi\ , \tag{34}$$

where $\hat{L}_p \in \left\{10^{-4}, 2 \times 10^{-6}, 10^{-6}, 5 \times 10^{-7}, 10^{-7}, 10^{-8}\right\}$ is an array of non-dimensional vascular permeabilities used in the numerical tests of [5]. In Fig. 5 we represent the capillary pressure as a function of arc-length for different vascular permeabilities. We can observe perfect agreement with the plots shown in [5] (not reported here). Moreover, we notice that for the lowest value of the vascular permeability $\hat{L}_p$, corresponding to an almost impermeable vessel, the computational model predicts a linearly decreasing pressure, in agreement with the Poiseuille equation that governs the flow. Conversely, for high permeability values there is a substantial deviation from the linear trend because the leakage dominates over the axial flow component.

In addition, a sensitivity analysis has been performed to investigate the influence of the relative position of the 3D and 1D grids. To this purpose, we simulated three different configurations in which the network vertexes coincide with particular points of the 3D mesh, as shown in Fig. 6. Numerical results of similar test cases, where the 1D mesh is slightly shifted to coincide with the location of Fig. 6 (top),

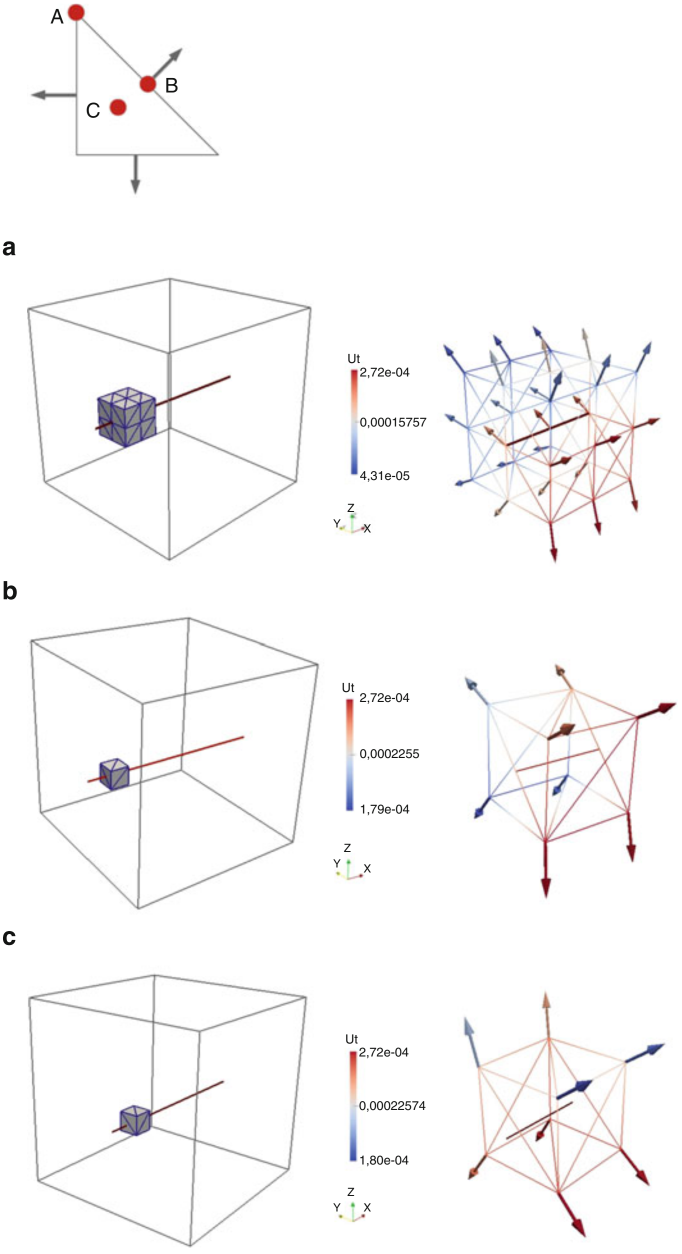


Fig. 6 (*Top panel*) Proposed configurations for 1D/3D mesh correlation analysis. The *red dots* A, B, C indicate three meaningful configurations w.r.t. the distribution of Raviart-Thomas dof (*arrows*). We show below the sensitivity analysis for the 1D/3D mesh coupling. Numerical solutions have been obtained with mesh size $h = 0.1$ and parameters $\kappa_t = \kappa_v = 1$, $R' = 1$, $Q = 10^{-4}$. For each of the three configurations A, B, C, we extract the smallest patch of elements intersected by the 1D mesh (truncated along the axial direction for visualization purposes). The *local* velocity field is also displayed together with its magnitude (color scale)

are given in Fig. 6 (bottom). These results suggest that when the 1D mesh is not aligned with edges or faces of the 3D one, the velocity field in the neighborhood of the vessel looks smooth and symmetric. A similar conclusion holds true also when the 1D branch lays on the 3D mesh faces. However, a problematic case is observed when the 1D mesh entirely coincides with edges of the 3D one. In this case, the velocity field looses cylindrical symmetry around the 1D capillary. A preliminary and heuristic interpretation of this behavior can be found observing that the Raviart-Thomas degrees of freedom are located on the element faces. For this reason, the 3D velocity field is not uniquely defined on tetrahedral element edges. When the 1D source term is exactly located on the element edges, there is an inconsistency in the approximation of the velocity field.

5.2 Coupled 3D-1D Problem on a Y-Shaped Bifurcation

We aim to validate the imposition of the mass conservation constraint at the junction. We observe that the conservation of total pressure at the junction $\mathbf{x}_M$,

$$p_v^0(\mathbf{x}_M) = p_v^1(\mathbf{x}_M) = p_v^2(\mathbf{x}_M) \equiv p_v(\mathbf{x}_M) \ , \tag{35}$$

is automatically ensured thanks to the use of continuous finite elements for the vessels pressure approximation. Conversely, in order to impose the mass conservation constraint

$$u_v^0(\mathbf{x}_M) = u_v^1(\mathbf{x}_M) + u_v^2(\mathbf{x}_M) \ , \tag{36}$$

we proceed as in (30). In practice, in the simple Y-shaped configuration, we first identify the FEM degrees of freedom (dofs) related to the same junction node. For those dofs we add in some specific entries of the problem matrix $+1$ for each inflow branch and -1 for each outflow branch.

5.2.1 Numerical Results

We apply again the *SuperLU* direct method to solve the linear system (31) and boundary conditions as in the second test-case, (33).

We notice that both vessels pressure and velocity confirm the expected behavior along the network: the former is continuous everywhere while the latter is split into two after the junction. We then conclude that conditions (35) and (36) are fulfilled also at the numerical level. Furthermore, Fig. 7 (bottom panel) confirms that, also in this case, the 3D/1D coupling behaves correctly.

Finally, we present in Fig. 8 a comparison against the pressure formulation of the 3D/1D coupled problem (6), previously proposed in [3, 4]. In that work, only the p_t and p_v variables were approximated, in particular using piecewise linear finite

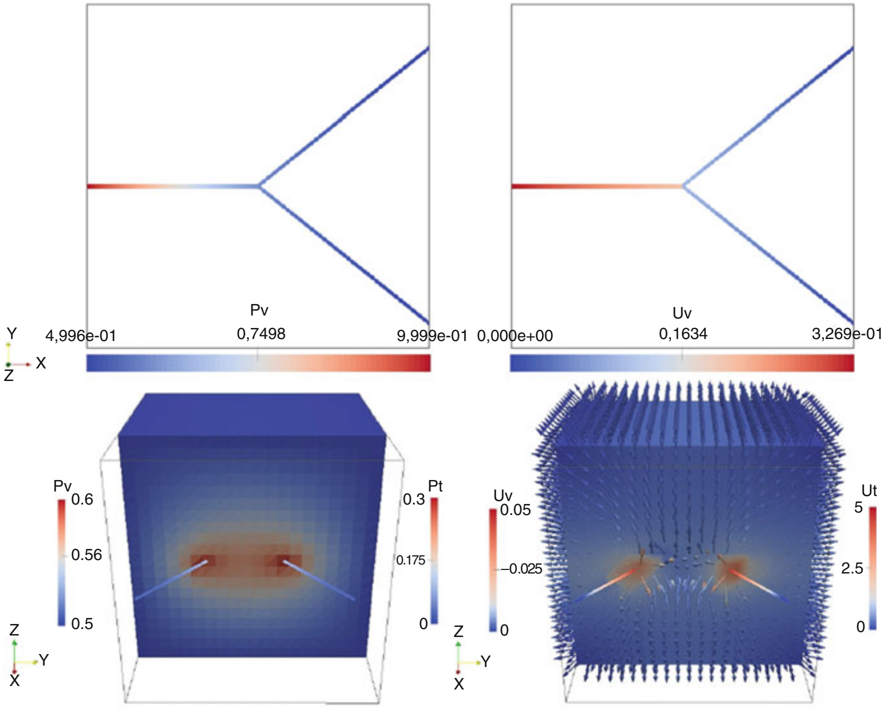


Fig. 7 (*Top*) Numerical solution of the 1D vessel problem, $(p_{v,h}, u_{v,h})$ obtained with $h = 0.05$ and unitary parameters $\kappa_v = 1$, $R' = 1$, $Q = 1$. As expected the pressure (*left*) is almost linearly decreasing between the imposed boundary values 1 and 0; velocity (*right*) is almost constant over each branch and it halves after the junction. (*Bottom*) Visualization of the 3D/1D coupled pressure and velocity fields

elements. The velocity field was reconstructed a-posteriori, after calculating the pressure gradients as constant vector functions over each element. By comparing the results of the *new* mixed-form model (left column) against those of the pressure-form model (right column) it appears, as expected, that we lose accuracy in 3D pressure approximation, because we use piecewise constant approximation instead of piecewise linears, but we visibly gain a better approximation of the 3D velocity field. This is ultimately a very important advantage. Indeed, following the work of [3, 16] we are planning to combine these simulations with mass transport problems for drug delivery through the microcirculation, where the velocity field is adopted to model advection.

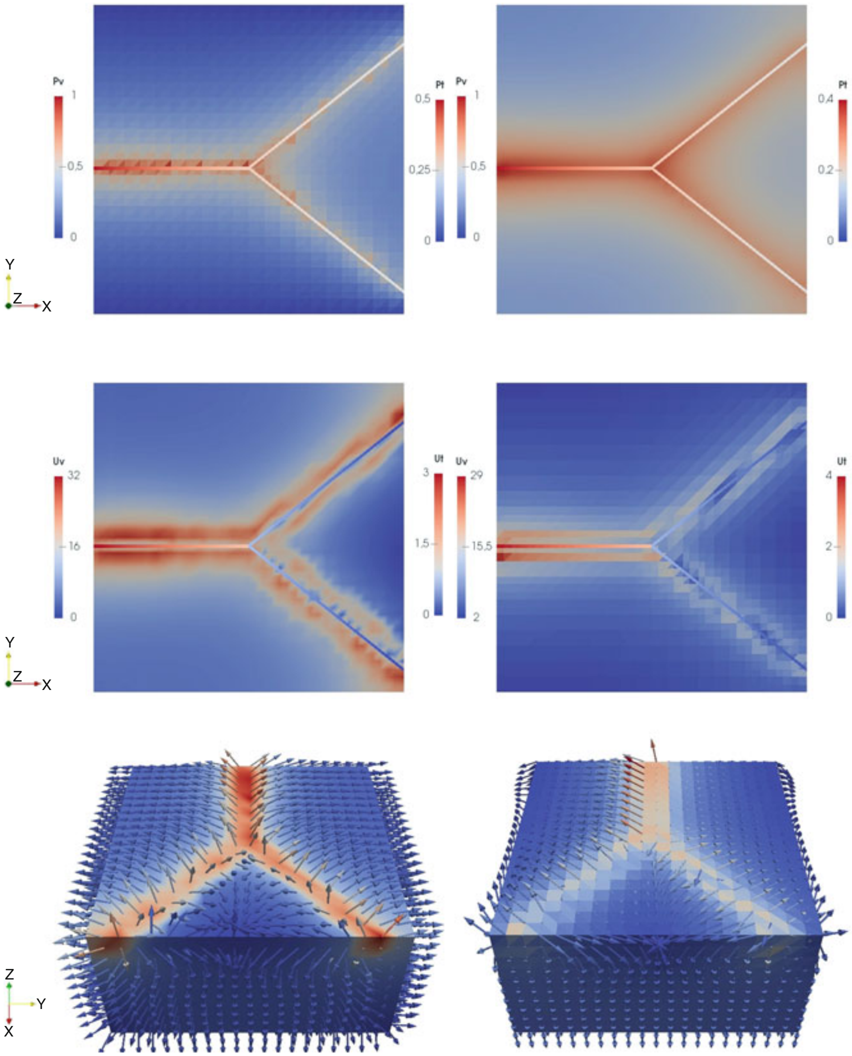


Fig. 8 Comparison between numerical predictions of pressure-form (*right column*) and mixed-form (*left column*) models within identical settings. In both cases we adopted $h = 0.05$ and dimensionless parameters $R' = 0.1$, $\kappa_t = 1$, $\kappa_v = 1$, $Q = 1$

6 Conclusions

We have proposed a mixed finite element formulation for coupled incompressible flow problems defined on a 1D domain embedded into a 3D porous medium. Because of the non standard coupling operators based on nonlocal restriction of the 3D solution to the 1D manifold, the definition and implementation of the method is challenging. We have implemented a non-standard finite element method into a C++ solver and the purpose of this work was to carefully validate it on two benchmark problems. The numerical solutions feature the expected behavior and confirm the correct functioning of the code. The mixed finite element approximation for 3D/1D coupled incompressible flow problems is a significant improvement with respect to the previously available solver based on the primal pressure formulation, because the velocity field is a variable of the problem and mass conservation constraints are directly enforced. Indeed, we are planning to exploit the better approximation properties of the velocity field, to combine the present solver with mass transport equations.

References

1. Blake, T., Gross, J.: Analysis of coupled intra- and extraluminal flows for single and multiple capillaries. Math. Biosci. **59**(2), 173–206 (1982)
2. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods, Springer Series in Computational Mathematics, vol. 15. Springer, New York (1991). doi:10.1007/978-1-4612-3172-1
3. Cattaneo, L., Zunino, P.: A computational model of drug delivery through microcirculation to compare different tumor treatments. Int. J. Numer. Methods Biomed. Eng. **30**(11), 1347–1371 (2014). doi:10.1002/cnm.2661
4. Cattaneo, L., Zunino, P.: Computational models for fluid exchange between microcirculation and tissue interstitium. Netw. Heterog. Media **9**(1), 135–159 (2014)
5. Chapman, S.J., Shipley, R.J., Jawad, R.: Multiscale modeling of fluid transport in tumors. Bull. Math. Biol. **70**(8), 2334–2357 (2008). doi:10.1007/s11538-008-9349-7
6. D'Angelo, C.: Finite element approximation of elliptic problems with dirac measure terms in weighted spaces: applications to one- and three-dimensional coupled problems. SIAM J. Numer. Anal. **50**(1), 194–215 (2012)
7. D'Angelo, C., Quarteroni, A.: On the coupling of 1d and 3d diffusion-reaction equations: application to tissue perfusion problems. Math. Models Methods Appl. Sci. **18**(8), 1481–1504 (2008)
8. Flieschman, G., Secomb, T., Gross, J.: Effect of extravascular pressure gradients on capillary fluid exchange. Math. Biosci. **81**(2), 145–164 (1986)
9. Fleischman, G., Secomb, T., Gross, J.: The interaction of extravascular pressure fields and fluid exchange in capillary networks. Math. Biosci. **82**(2), 141–151 (1986)
10. Formaggia, L., Gerbeau, J., Nobile, F., Quarteroni, A.: On the coupling of 3d and 1d Navier-stokes equations for flow problems in compliant vessels. Comput. Methods Appl. Mech. Eng. **191**(6–7), 561–582 (2001). doi:10.1016/S0045-7825(01)00302-4
11. Glowinski, R., Pan, T.W., Periaux, J.: A fictitious domain method for Dirichlet problem and applications. Comput. Methods Appl. Mech. Eng. **111**(3–4), 283–303 (1994)
12. Glowinski, R., Pan, T.W., Hesla, T., Joseph, D.: A distributed lagrange multiplier/fictitious domain method for particulate flows. Int. J. Multiph. Flow **25**(5), 755–794 (1999)
13. Hicks, K., Pruijn, F., Secomb, T., Hay, M., Hsu, R., Brown, J., Denny, W., Dewhirst, M., Wilson, W.: Use of three-dimensional tissue cultures to model extravascular transport and predict in vivo activity of hypoxia-targeted anticancer drugs. J. Natl. Cancer Inst. **98**(16), 1118–1128 (2006)
14. Lesinigo, M., D'Angelo, C., Quarteroni, A.: A multiscale Darcy-Brinkman model for fluid flow in fractured porous media. Numer. Math. **117**(4), 717–752 (2011)

15. Mittal, R., Iaccarino, G.: Immersed boundary methods. Annu. Rev. Fluid Mech. **37**, 239–261 (2005)
16. Nabil, M., Decuzzi, P., Zunino, P.: Modelling mass and heat transfer in nano-based cancer hyperthermia. R. Soc. Open Sci. **2**(10), 150447, 1–17 (2015). doi:10.1098/rsos.150447
17. Peskin, C.S.: The immersed boundary method. Acta Numer. **11**, 479–517 (2002). doi:10.1017/S0962492902000077
18. Quarteroni, A., Formaggia, L., Veneziani, A.: Cardiovascular mathematics: modeling and simulation of the circulatory system. Model. Simul. Appl. **1**, 1–512 (2009)
19. Secomb, T., Hsu, R., Braun, R., Ross, J., Gross, J., Dewhirst, M.: Theoretical simulation of oxygen transport to tumors by three-dimensional networks of microvessels. Adv. Exp. Med. Biol. **454**, 629–634 (1998)
20. Secomb, T., Hsu, R., Park, E., Dewhirst, M.: Green's function methods for analysis of oxygen delivery to tissue by microvascular networks. Ann. Biomed. Eng. **32**(11), 1519–1529 (2004)
21. Yu, Y., Baek, H., Karniadakis, G.: Generalized fictitious methods for fluid-structure interactions: analysis and simulations. J. Comput. Phys. **245**, 317–346 (2013)
22. Zhang, L., Gerstenberger, A., Wang, X., Liu, W.K.: Immersed finite element method. Comput. Methods Appl. Mech. Eng. **193**(21–22), 2051–2067 (2004). doi:10.1016/j.cma.2003.12.044

On a PDE-Constrained Optimization Approach for Flow Simulations in Fractured Media

Sandra Pieraccini and Stefano Scialò

Abstract In the present work we consider the problem of performing underground flow simulations in fractured media, following the Discrete Fracture Network (DFN) model. We will focus on a quite recent approach to the problem, based on a PDE-constrained optimization formulation, which allows for the use of totally non-conforming meshes on the network. In this way arbitrarily complex DFNs can be effectively tackled, without requiring any modification of the geometry of the network. Extended numerical simulations are reported demonstrating the performances of the proposed method, and highlighting its robustness in handling networks with hard-to-mesh configurations, such as extremely narrow angles between intersecting fractures. The problem of advection-diffusion of pollutant species in networks of fractures is also addressed in a time-dependent framework, using the optimization-based approach both to derive the Darcy velocity and to solve the transport problem at each time frame.

1 Introduction

Several critical and up-to-date applications involving underground exploitation, such as geothermal applications, enhanced oil & gas production, geological storage, require effective tools for performing underground flow simulations, possibly in very large scale domains representing basins.

Within this framework, a possible approach for modelling the underground is given by the Discrete Fracture Network (DFN) model [1, 17, 20], in which the fractures in the rock matrix are represented as planar polygons intersecting each other along segments called *traces*. The surrounding rock matrix is assumed here to be impervious, so that no flux exchange occurs among the fractures and the surrounding rock matrix. Nevertheless, flux exchanges occur among fractures through the traces. The quantity under investigation is the hydraulic head H, which

---

S. Pieraccini • S. Scialò (✉)
Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy
e-mail: sandra.pieraccini@polito.it; stefano.scialo@polito.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_2

is related to the fluid pressure p, being defined as $H = p/(\rho g) + z$, where g is the gravitational acceleration, ρ the fluid density and z the elevation. The flow on each fracture is modelled by means of the Darcy law, and suitable matching conditions are imposed at fracture intersections, over the whole network, in order to ensure head continuity and flux balance.

The major complexity in approaching the problem of the effective simulation of the flow in large networks of fractures lies in the geometrical size and complexity of the computational domain, characterized by a large number of fractures, with mutually intersecting traces that might form extremely narrow angles, or might be extremely close to each other. These hard-to-mesh configurations make the use of standard approaches based on conforming grids infeasible, and require targeted methods in order to ease the meshing process. Solutions proposed in [19, 34–37] make use of the mortar method, sometimes in conjunction with modifications of the geometry. In [18, 31, 32] problem complexity is tackled resorting to the resolution of lower-dimensional problems; in [28] an algorithm is designed, suitable for stochastic analysis in DFNs, capable of detecting and rejecting DFN configurations too complex to be meshed. Other techniques to solve complex flow configurations in DFNs can be found in [21, 22, 26, 29, 30].

In rather recent papers [11–13, 15] a new approach has been proposed in order to tackle any kind of geometrical configuration without requiring any form of mesh conformity at fracture intersections. The method has a strongly parallel nature and, thanks to the nonconformity, allows for an independent mesh generation process on each fracture of the DFN [14], thus overcoming the complexities encountered by classical approaches. The method can be used with standard finite element discretizations, but it is especially designed in order to take advantage of non-standard discretization techniques, such as the XFEM [16] and the VEM [8].

In the present work we briefly sketch a formulation of the flow problem in networks of fractures (Sect. 2) and we shortly recall the mathematical foundations of the method and its discrete formulation in Sect. 3. New and extensive numerical results are presented in Sect. 4 on complex DFN configurations that can be considered realistic from the point of view of the geometrical complexities. Both the problem of the computation of the steady-state distribution of the hydraulic head and the problem of the transient response in advection-diffusion phenomena are addressed.

2 Problem Formulation

In this section we briefly recall the problem formulation on the network. We refer the reader to [11, 13] for more details.

Each fracture in the network is modelled as a planar polygon, with arbitrary dimensions and orientation in space. Let F_i denote the i-th fracture, with $i \in \mathscr{I}$. The fractures may intersect each other along a line called trace. Let S_m denote the

m-th trace, with $m \in \mathscr{M}$. In the sequel we assume that each trace is generated by exactly two fractures.

In order to relate the trace index m to the fracture indices i and j generating S_m, we set $I_m = (i,j)$, being the couple (i,j) ordered following the convention that $i < j$.

The DFN Ω is therefore defined as the set

$$\Omega = \cup_{i \in \mathscr{I}} F_i.$$

Each fracture boundary ∂F_i is (possibly) divided in a Dirichlet part Γ_i^D and in a Neumann part Γ_i^N. We correspondingly introduce the following sets of fracture boundaries:

$$\partial\Omega = \cup_{i \in \mathscr{I}} \partial F_i, \quad \Gamma^D = \cup_{i \in \mathscr{I}} \Gamma_i^D, \quad \Gamma^N = \cup_{i \in \mathscr{I}} \Gamma_i^N.$$

Notice that we require Γ^D to be non-empty; on the other hand, some of the sets Γ_i^D are allowed to be empty. For further reference, we introduce the following sets: we define

$$\mathscr{S} = \cup_{m \in \mathscr{M}} S_m$$

and $\forall i \in \mathscr{I}$ we denote by $\mathscr{S}_i \subset \mathscr{S}$ the subset of traces belonging to fracture F_i.

Let us introduce $\forall i \in \mathscr{I}$ the spaces

$$V_i = \mathrm{H}_0^1(F_i) = \left\{ v \in \mathrm{H}^1(F_i) : v_{|_{\Gamma_i^D}} = 0 \right\}$$

with dual space V_i', and, if $\Gamma_i^D \neq \emptyset$, let us introduce

$$V_i^D = \mathrm{H}_\mathrm{D}^1(F_i) = \left\{ v \in \mathrm{H}^1(F_i) : v_{|_{\Gamma_i^D}} = H_i^D \right\},$$

where H_i^D is the Dirichlet condition imposed on Γ_i^D. In the sequel, we adopt the convention that for functions globally defined on the whole network, a i subscript denotes the restriction to F_i.

Let $\mathbf{K}_i(x_i)$ be, for all $i \in \mathscr{I}$, a symmetric and uniformly positive definite tensor describing the fracture transmissivity, where x_i refers to the local space coordinate system on F_i. The hydraulic head on each fracture is obtained as follows: find $H_i \in V_i^D$ such that $\forall v \in V_i$

$$\int_{F_i} \mathbf{K}_i \nabla H_i \nabla v \mathrm{d}\Omega = \int_{F_i} q_i v \mathrm{d}\Omega + \int_{\Gamma_i^N} H_i^N v_{|\Gamma_i^N} \mathrm{d}\Gamma + \sum_{S \in \mathscr{S}_i} \int_S \left[\!\!\left[\frac{\partial H_i}{\partial \nu_S^i} \right]\!\!\right]_S v_{|S} \mathrm{d}\Gamma, \quad (1)$$

where $H_i^N \in \mathrm{H}^{-\frac{1}{2}}(\Gamma_i^N)$ is the Neumann boundary condition imposed on Γ_i^N. The symbol $\frac{\partial H_i}{\partial \nu_S^i} = (n_S^i)^T \mathbf{K}_i \nabla H_i$ denotes the outward co-normal derivative of the hydraulic head along the (fixed) unit vector n_S^i normal to trace S, and $\left[\!\left[\frac{\partial H_i}{\partial \nu_S^i}\right]\!\right]_S$ denotes its jump along n_S^i. Furthermore, $q_i \in \mathrm{L}^2(F_i)$ is a source term on F_i. Note that the last term in (1) represents the net flow entering/exiting the fracture through its traces.

In order to ensure continuity of H and flux balance along all the network, the following matching conditions are imposed at all traces: $\forall m \in \mathscr{M}$, with $i,j \in I_m$,

$$H_{i|S_m} - H_{j|S_m} = 0, \tag{2}$$

$$\left[\!\left[\frac{\partial H_i}{\partial \nu_{S_m}^i}\right]\!\right]_{S_m} + \left[\!\left[\frac{\partial H_j}{\partial \nu_{S_m}^j}\right]\!\right]_{S_m} = 0. \tag{3}$$

3 A PDE-Constrained Optimization Approach

We describe here an approach to the solution of problem (1)–(3) first conceived in [11, 12] and further developed in [8, 13]. In particular, we follow the approach of the latter reference. In order to ease notation, in the following we assume that the traces are disjoint; this is a simplifying assumption which can be easily removed, see [11, Remark 2.1].

Let $\mathscr{U}_i := \prod_{S \in \mathscr{S}_i} \mathrm{H}^{-\frac{1}{2}}(S)$, $\mathscr{U} := \prod_{i \in \mathscr{I}} \mathscr{U}_i$. For each $m \in \mathscr{M}$, with $I_m = (i,j)$, we define the quantities

$$U_i^m := \left[\!\left[\frac{\partial H_i}{\partial \nu_{S_m}^i}\right]\!\right]_{S_m} + \alpha H_{i|S_m}, \qquad U_j^m := \left[\!\left[\frac{\partial H_j}{\partial \nu_{S_m}^j}\right]\!\right]_{S_m} + \alpha H_{j|S_m}, \tag{4}$$

where $\alpha > 0$ is a fixed parameter, which will play the role of control variables. Next, we build the following tuples of control variables:

$$U_i = \prod_{S \in \mathscr{S}_i} U_i^S \in \mathscr{U}_i, \qquad U = \prod_{i \in \mathscr{I}} U_i \in \mathscr{U}\,.$$

The exact fulfilment of Eqs. (2) and (3) is replaced by the minimization of the functional

$$J(H,U) = \sum_{m \in \mathscr{M}} \Bigg(\left\| H_{i|S_m} - H_{j|S_m} \right\|^2_{\mathrm{H}^{\frac{1}{2}}(S_m)} + \tag{5}$$

$$\left\| U_i^{S_m} + U_j^{S_m} - \alpha \left(H_{i|S_m} + H_{j|S_m} \right) \right\|^2_{\mathrm{H}^{-\frac{1}{2}}(S_m)} \Bigg),$$

with i,j such that $I_m = (i,j)$, constrained by Eq. (1), which—taking into account (4)—rewrites

$$\int_{F_i} \mathbf{K}_i \nabla H_i \nabla v \mathrm{d}\Omega + \alpha \sum_{m\in\mathcal{M}} \int_{S_m} H_{i|S_m} v_{|S_m} \mathrm{d}\gamma = \int_{F_i} q_i v \mathrm{d}\Omega + \int_{\Gamma_{iN}} H_i^N v_{|\Gamma_{iN}} \mathrm{d}\gamma + \sum_{m\in\mathcal{M}} \int_{S_m} U_i^m v_{|S_m} \mathrm{d}\gamma \tag{6}$$

$\forall v \in V_i$, $\forall i \in \mathscr{I}$. The problem

$$\begin{aligned} &\min \; J(H,U) \\ &\text{subject to} \quad (6) \end{aligned} \tag{7}$$

has a unique minimizer which corresponds to the solution of (1)–(3), see [11, 13].

By using a finite element based discretization and following a "first discretize, then optimize" approach, a finite dimensional counterpart of problem (7) is obtained. To this aim, we introduce on each fracture a triangular mesh, independently of the position of the other fractures and of fracture intersections; we also introduce a space discretization on the traces. Since the hydraulic head may be nonsmooth along the traces, a special care is needed for an accurate reproduction of the solution near the traces. This has been accomplished, in conjunction with this approach, following several strategies. In [11–13] the Extended Finite Element method (XFEM) [7, 25] has been used; in [8] the original triangular mesh has been transformed in a polygonal mesh, using the new Virtual Element Method [2–6]. The description which follows does not rely on a specific space discretization.

We introduce:

- on each fracture F_i, $i \in \mathscr{I}$, a (possibly non-standard) finite element basis $\{\varphi_{i,k}\}$, $k = 1, \dots, N_i$;
- on each trace S_m, $m \in \mathcal{M}$, the set $\{\psi_{i,k}^m\}$, $k = 1, \dots, N_i^m$ of 1D linear finite element basis functions.

In the sequel we will use small case letters to denote the discrete variables: in particular, h_i will denote the discrete hydraulic head on F_i and u_i^m will denote the discrete control variable on the trace S_m of fracture F_i:

$$h_i = \sum_{k=1}^{N_i} h_{i,k} \varphi_{i,k}, \qquad u_i^m = \sum_{k=1}^{N_i^m} u_{i,k}^m \psi_{i,k}^m. \tag{8}$$

With a notation overload, h_i and u_i^m will also denote the corresponding vectors of DOFs, the difference being always clear from the context.

In view of writing the discrete counterpart of problem (7), we introduce for each fixed $i \in \mathscr{I}$ the vector $u_i \in \mathbb{R}^{N_{\mathscr{S}_i}}$ obtained by concatenating (column-wise) vectors u_i^m corresponding to traces $S_m \in \mathscr{S}_i$, thus $N_{\mathscr{S}_i} = \sum_{S_m \in \mathscr{S}_i} N_i^m$. We further concatenate column-wise vectors h_i, with $i \in \mathscr{I}$, obtaining a vector $h \in \mathbb{R}^{N^F}$, $N^F = \sum_{i\in\mathscr{I}} N_i$, and vectors u_i obtaining a vector $u \in \mathbb{R}^{N^T}$, $N^T = \sum_{i\in\mathscr{I}} N_{\mathscr{S}_i}$.

The constraint Eqs. (6) on each fracture rewrite as

$$A_i h_i = \tilde{q}_i + \mathscr{B}_i u_i, \tag{9}$$

where $A_i \in \mathbb{R}^{N_i\times N_i}$ is the stiffness matrix, matrix $\mathscr{B}_i \in \mathbb{R}^{N_i\times N_{\mathscr{S}_i}}$ collects the integrals of functions $\varphi_{i,k}$, $k = 1,\ldots,N_i$, against functions $\{\psi^m_{i,k}\}$, $k = 1,\ldots,N^m_i$, $S_m \in \mathscr{S}_i$, and the vector $\tilde{q}_i \in \mathbb{R}^{N_i}$ accounts for the source term and the boundary conditions.

Now we turn our attention to the discrete counterpart of (5). While referring the reader to [13] for all the details, we give here a brief sketch of the discrete formulation.

In Eq. (5), we compute J by means of $\mathrm{L}^2(S_m)$ norms and plug into (5) definition (8). The algebraic form obtained for (7) is as follows:

$$\min \quad J(h,u) := \frac{1}{2}\left(h^T G^h h - \alpha h^T B u - \alpha u^T B^T h + u^T G^u u\right) \tag{10}$$

$$\text{s.t.} \quad Ah = \mathscr{B}u + q, \tag{11}$$

where $A \in \mathbb{R}^{N^F\times N^F}$ is a block diagonal matrix built from the local stiffness matrices A_i; $G^h \in \mathbb{R}^{N^F\times N^F}$ and $G^u \in \mathbb{R}^{N^T\times N^T}$ are sparse matrices defined by suitably assembling the local blocks defined on the fractures, and collect integrals of the basis functions for h and u, respectively, on the traces, i.e. integrals of the kind:

$$\int_{S_m} \varphi_{i,k|S_m}\varphi_{j,\ell|S_m}, \qquad \int_{S_m} \psi^m_{i,k}\psi^m_{j,\ell}; \tag{12}$$

for either $i = j$ or $(i,j) = I_m$; $B \in \mathbb{R}^{N^F\times N^T}$, $\mathscr{B} \in \mathbb{R}^{N^F\times N^T}$ are again sparse matrices defined by suitably assembling the local blocks defined on the fractures, and collect integrals of functions $\varphi_{i,k}$ against functions $\psi^m_{j,k}$.

Notice that despite Eqs. (10)–(11) are defined on the whole DFN, the computations can be actually decoupled in *local* fracture-oriented computations, see [14]. Linearity of the constraints (11) easily allows to eliminate h from J, thus obtaining the following equivalent unconstrained minimization problem:

$$\begin{aligned}\min \hat{J}(u) &:= \frac{1}{2}u^T(\mathscr{B}^T A^{-T} G^h A^{-1}\mathscr{B} + G^u - \alpha\mathscr{B}^T A^{-T} B - \alpha B^T A^{-1}\mathscr{B})u \\ &\quad + q^T A^{-T}(G^h A^{-1}\mathscr{B} - \alpha B)u \\ &:= \frac{1}{2}u^T \hat{G} u + \hat{q}^T u. \end{aligned} \tag{13}$$

3.1 Discretization Strategies

The proposed optimization approach does not require any kind of conformity along the traces for the meshes on the intersecting fractures, as only integrals of the kind of (12) need to be computed, that involve the discretization on different fractures. Typically, the hydraulic head displays a non-smooth behaviour on the fractures, as jumps in the first order derivatives may occur across the traces, corresponding to the flux entering or leaving the fracture through its traces. This non-smooth behaviour can not be correctly reproduced with standard finite elements on nonconforming meshes, and for this reason unconventional discretization techniques might be preferred. A possible strategy consists in using the flexibility of the VEM in handling polygonal meshes to generate a mesh on each fracture that is conforming to the traces on the fracture but is still nonconforming to the mesh on the intersecting fractures. The resulting meshing process is independent on each fracture and is performed by simply splitting the triangular elements of the nonconforming mesh crossed by the traces in sub-polygons not crossed by the traces. This idea is at the basis of the approaches described in [8–10]. Another strategy, adopted for the simulations in the present work, consists in the use of the XFEM, thus adding to the discrete space suitable basis functions reproducing the irregular behaviour of the solution. The use of the XFEM in the context of DFN simulations in conjunction with the optimization approach was proposed since the earliest works on the subject [11–13], and is also adopted in different frameworks, as for example in [26, 27]. Some numerical issues need to be addressed for efficient use of the XFEM in large scale DFN simulations with the proposed approach, mainly concerning prevention of ill conditioning and the imposition of boundary conditions. Although these topics are extensively discussed in the literature (see e.g. [24, 33]), targeted solutions are mandatory in the present context. The use of XFEM might be source of ill conditioning due to the presence of redundant, linearly (or nearly linearly) dependent basis functions in the discrete space, as a consequence of the additional basis functions introduced to catch the behaviour of the solution near the traces. When parallel or nearly parallel traces intersect the same mesh element, some of the resulting basis functions might be redundant. Due to the complexity of the geometry, a preliminary detection of redundant basis functions is not a viable option, and other strategies need to be performed, such as a fracture-local rank revealing QR factorization [16]. Also the imposition of Dirichlet boundary conditions can be performed weakly through the functional, following the approach in [16].

4 Numerical Results

In this Section we report some numerical results aimed at showing the viability, the performances, and the robustness of the proposed strategy for the simulation of the flow in complex DFNs. A first set of simulations deals with the computation of the hydraulic head distribution in realistic networks of varying complexity and number of fractures, followed by a second set of numerical results reporting some preliminary results on the resolution of diffusion-advection problems in DFNs with the optimization-based approach.

4.1 Flow in Complex DFNs

In this Section we deal with the steady-state solution on DFNs with realistic geometrical configurations, exhibiting challenging critical features such as very narrow angles between intersecting traces, or very close and almost parallel traces lying in the same fracture. These characteristics represent the major causes of failure for simulations based on standard approaches relying on finite elements on conforming grids, whereas they are easily tackled by the optimization based method considered here. In order to prove the robustness of the method in dealing with such hard-to-mesh features, we will compare the performances of the method on networks with severe geometrical configurations with the results on less complex networks.

A pool of four networks is considered here, with increasing number of fractures. In Table 1 we report information concerning the network sizes. Namely, in the table $\mathscr{I}^\#$ and $\mathscr{M}^\#$ represent the number of fractures and traces in the network, respectively, and L is the cumulative length of all the traces. The two intermediate size networks (*709fract* and *909fract*) are stochastically generated, whereas the *609fract* and *959fract* networks are obtained starting from stochastically generated networks, removing some fractures in order to eliminate the mentioned hard-to-mesh geometrical features, thus obtaining simpler geometrical configurations.

The result of this process is shown in Figs. 1 and 2, in which we report the distribution of the hard-to-mesh features for the considered networks: the distribution of angles formed by couples of intersecting traces in the same fracture is shown in Fig. 1, highlighting that the two intermediate size networks *709fract* and *909fract* have angles approaching zero (the smallest value is less than 0.3° for both DFNs), while the *609fract* and the *959fract* DFNs do not display angles smaller than 10°. In Fig. 2 the distribution of distances between couples of non intersecting traces in the same fracture is reported, underlining again that the *709fract* and *909fract* DFNs have traces very close to each other, which are not present in the other two networks.

Table 1 DFN labels and data concerning network size

DFN label	$\mathscr{I}^\#$	$\mathscr{M}^\#$	L
609fract	609	1690	$2.5922e+4$
709fract	709	3939	$6.6206e+4$
909fract	909	7084	$1.1798e+5$
959fract	959	4911	$7.5822e+4$

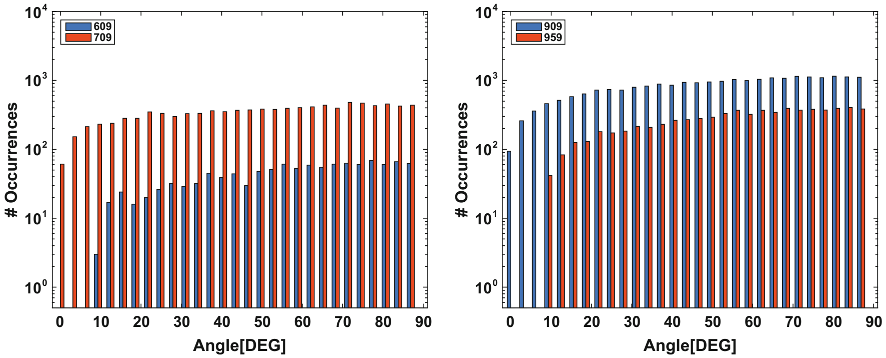


Fig. 1 Distribution of the angle formed by couples of intersecting traces in a fracture: *Left*: *609fract* and *709fract* DFNs; *Right*: *909fract* and *959fract* DFNs

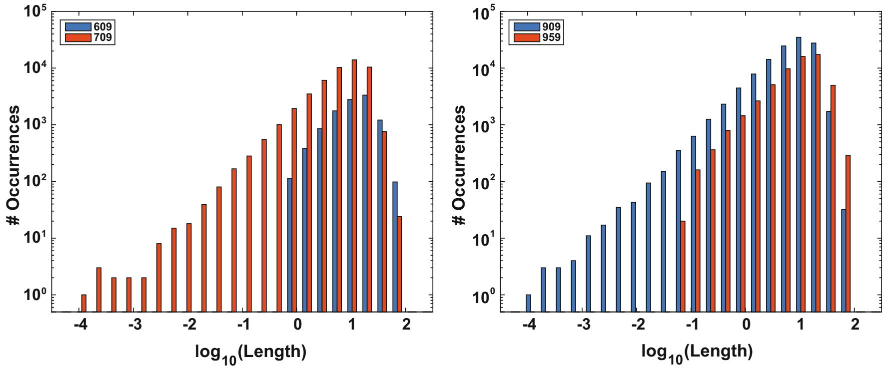


Fig. 2 Distribution of the distance between couples of non intersecting traces in a fracture: *Left*: *609fract* and *709fract* DFNs; *Right*: *909fract* and *959fract* DFNs

A non uniform fracture transmissivity field is prescribed, and the fracture transmissivity on F_i is given by:

$$\mathbf{K}_i = \begin{cases} 10^{-4}\left(1 + c_{i,1}\mathrm{e}^{\frac{-4}{1+16c_{i,2}\left((\tilde{x}-1)^2+(\tilde{y}-1)^2\right)}}\right) & \text{if } i \text{ is odd} \\ 10^{-4}\left(1 + c_{i,1}\mathrm{e}^{-2c_{i,2}\left((\tilde{x}-1)^2+(\tilde{y}-1)^2\right)}\right) & \text{if } i \text{ is even} \end{cases} \tag{14}$$

where constants $c_{i,1} \in \{10^{-2}, 10^{-1}, 1\}$ and $c_{i,2} \in \{0.1, 0.5, 1, 2\}$ are randomly chosen on each fracture and $\tilde{x} = \dfrac{2x}{\ell_x}$ and $\tilde{y} = \dfrac{2y}{\ell_y}$, being ℓ_x, ℓ_y characteristic dimensions of the fracture in the x and y-direction, respectively, in the local reference system.

On all the considered networks, a source and a sink fracture are placed at the top and the bottom of the domain of interest, respectively, carrying Dirichlet boundary

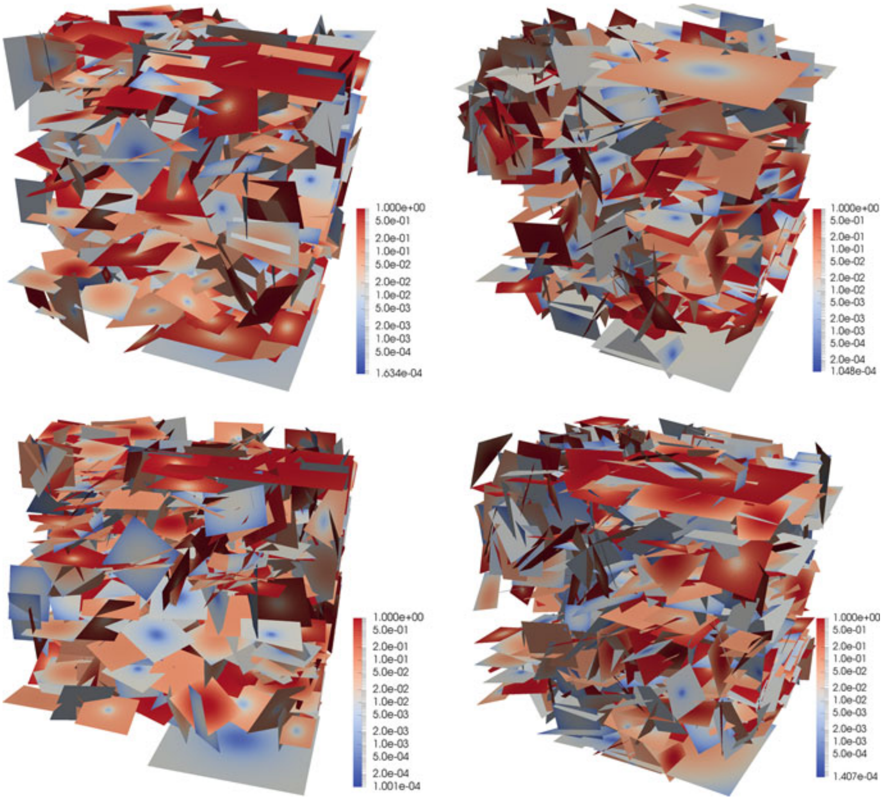


Fig. 3 DFN geometry and fracture transmissivity distribution (log scale): left to right, top to bottom: *609fract*, *709fract*, *909fract*, *959fract*

conditions, prescribing a value of $H^D = 1000$ on one edge of the source fracture and $H^D = 0$ on one edge of the sink fracture. All other fracture edges are insulated. The DFN geometries, along with the distribution of the fracture transmissivities, are shown in Fig. 3 for the considered networks.

Simulations are performed on three different grids, characterized by the mesh parameter $\delta \in \{7, 2, 0.5\}$, representing the maximum element area of the mesh. The total number of degrees of freedom (DOFs) for each DFN and for each grid are reported in Table 2. The saddle point problem is solved through the parallel iterative solver described in [14], implementing the preconditioned conjugate gradient method. The preconditioner P is a three-diagonal symmetric positive definite matrix representing a very rough approximation of the matrix $\hat{G}$ in (13), and has a block diagonal structure, with blocks related to the fractures in the DFN, [16]. The performances of the preconditioner are shown for the *709fract* network in Table 3, which reports the number of iterations required by the iterative solver to reach a relative residual (namely, relative to the norm of the right hand side of the linear system) smaller than a prescribed value σ, both in the preconditioned (*prec*) and in the not preconditioned (*not prec*) case. It can be seen that the

Table 2 Number of degrees of freedom for each DFN and mesh parameter

DFN	$\delta = 7$	$\delta = 2$	$\delta = 0.5$
609fract	230,723	552,615	1,681,341
709fract	442,073	941,475	2,496,875
909fract	733,486	1,498,025	3,751,857
959fract	541,888	1,175,673	3,215,716

Table 3 Number of iterations vs σ

	$\sigma = 10^{-6}$		$\sigma = 10^{-7}$		$\sigma = 10^{-8}$	
δ	Not prec	Prec	Not prec	Prec	Not prec	Prec
7	34,774	14,692	241,974	46,171	388,523	75,480
2	26,204	14,913	106,250	32,097	347,048	81,459
0.5	23,149	14,896	87,542	37,983	324,108	95,563

Comparison of the preconditioned and not preconditioned case on the *709fract* DFN

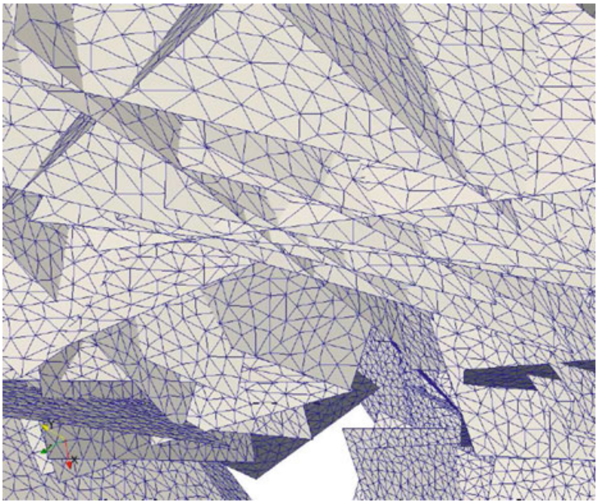


Fig. 4 Detail of the computational mesh with $\delta = 2$ for the *709fract* DFN

preconditioner provides a reduction in the number of iterations with respect to the not preconditioned case up to a factor of about 5 for the coarsest mesh and the smallest relative residual. It is to remark that, thanks to its block-diagonal structure, the preconditioner P can be applied inexpensively also in the parallel version of the algorithm.

A detail of the computational mesh with $\delta = 2$ is shown in Fig. 4 for the *709fract* DFN: it can be noticed that mesh element edges are arbitrarily placed with respect to the intersections between fractures, as a result of the meshing process that is performed on each fracture independently of the other fractures and of the traces. The solution obtained for the *709fract* and *959fract* networks is reported, as an example, in Figs. 5 and 6, respectively, on the intermediate size mesh ($\delta = 2$). In

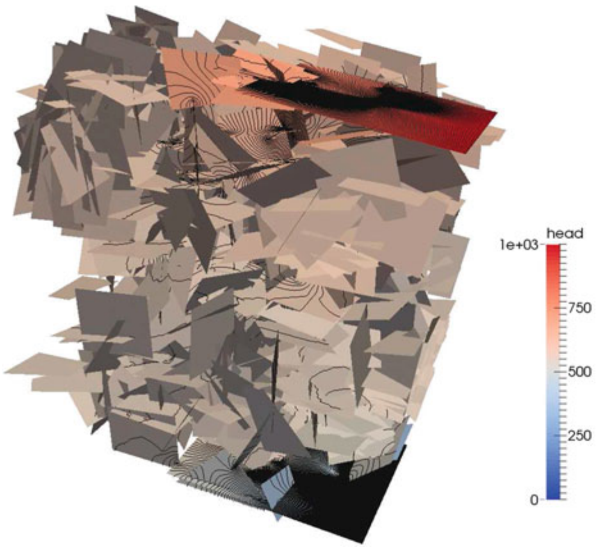


Fig. 5 Solution of the *709fract* DFN on the mesh with $\delta = 2$ with iso-h lines

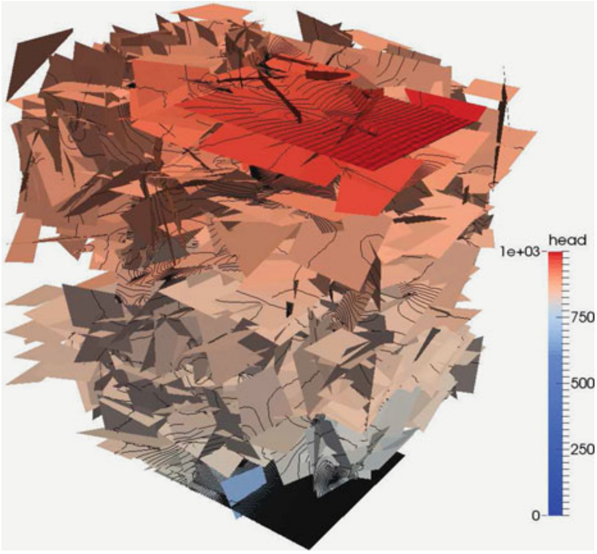


Fig. 6 Solution of the *959fract* DFN on the mesh with $\delta = 2$ with iso-h lines

these figures the colouring is proportional to the hydraulic head, and iso-h lines are also shown in order to highlight gradients in the solution.

In order to assess the robustness of the method in tackling critical geometrical configurations, typical of realistic DFNs, we compare the performances of the method in solving the four considered networks, recalling that DFNs *709fract* and *909fract* have hard-to-mesh features that are not present in the other two networks. Performances will be evaluated in terms of the number of iterations required by the iterative solver to reach a given relative residual and in terms of three error indicators used to determine the quality of the obtained solution. The error indicators are $\varDelta_{\text{head}}$, measuring the mismatch in hydraulic head continuity across the traces, $\varDelta_{\text{flux}}$, measuring the flux unbalance across the traces, and $\varDelta_{\text{cons}}$, representing the error in the conservation of the flux in the whole network. They are defined as:

$$\varDelta_{\text{head}} = \frac{\sqrt{\sum_{m\in\mathscr{M}} \left\| h_{i|_{S_m}} - h_{j|_{S_m}} \right\|^2}}{h_{\max} L},$$

$$\varDelta_{\text{flux}} = \frac{\sqrt{\sum_{m\in\mathscr{M}} \left\| u_i^m + u_j^m - \alpha_m \left(h_{i|_{S_m}} + h_{j|_{S_m}} \right) \right\|^2}}{\bar{\varPhi} L},$$

and

$$\varDelta_{\text{cons}} = \frac{\left| \sum_{i\in\mathscr{I}_{\text{in}}\cup\mathscr{I}_{\text{out}}} \sum_{S_m\in\mathscr{S}_i} \int_{S_m} (u_i^m - \alpha_m h_{i|S_m}) \right|}{\bar{\varPhi}},$$

where L is the cumulative trace length (see Table 1), $h_{\max}$ is the maximum of the hydraulic head in the network (that for the considered cases coincides with the Dirichlet condition on the source fracture) and $\bar{\varPhi}$ is the total flux flowing through the network. According to the imposed boundary conditions, $\bar{\varPhi}$ can be computed as:

$$\bar{\varPhi} = \frac{1}{2} \left(\sum_{i\in\mathscr{I}_{\text{in}}} \sum_{S_m\in\mathscr{S}_i} \int_{S_m} (u_i^m - \alpha_m h_{i|S_m}) - \sum_{i\in\mathscr{I}_{\text{out}}} \sum_{S_m\in\mathscr{S}_i} \int_{S_m} (u_i^m - \alpha_m h_{i|S_m}) \right),$$

being $\mathscr{I}_{\text{in}}$ and $\mathscr{I}_{\text{out}}$ the index sets corresponding to source and sink fractures, respectively. Following from the above definitions, the error indicators are relative quantities: $\varDelta_{\text{head}}$ expresses a measure of the error per unitary value of the source hydraulic head and per unitary trace length; $\varDelta_{\text{flux}}$ a measure of the error per unitary total flux and per unitary trace length; $\varDelta_{\text{cons}}$ a measure of the error per unitary total flux, this last quantity being independent of the number of traces in the network. These definitions as relative quantities allow the use of these error indicators to compare the quality of the solution obtained on DFNs with different geometries and hydraulic properties. Table 4 reports, for all the considered networks and grids, the number of iterations required by the preconditioned conjugate gradient method to reach values of the relative residual lower than three prescribed values ($\sigma = 10^{-6}, 10^{-7}, 10^{-8}$), whereas values for the three error indicators described

Table 4 Number of iterations vs σ for all the considered networks and meshsizes

609fract				*909fract*			
δ	$\sigma = 10^{-6}$	$\sigma = 10^{-7}$	$\sigma = 10^{-8}$	δ	$\sigma = 10^{-6}$	$\sigma = 10^{-7}$	$\sigma = 10^{-8}$
7	14,871	34,623	50,126	7.0	11,075	31,241	120,228
2	13,993	21,601	60,561	2.0	11,599	26,854	169,311
0.5	16,173	25,135	69,331	0.5	11,741	29,241	70,143
709fract				*959fract*			
δ	$\sigma = 10^{-6}$	$\sigma = 10^{-7}$	$\sigma = 10^{-8}$	δ	$\sigma = 10^{-6}$	$\sigma = 10^{-7}$	$\sigma = 10^{-8}$
7	14,692	46,171	75,480	7.0	24,043	42,980	64,869
2	14,913	32,097	81,459	2.0	26,483	47,671	66,662
0.5	14,896	37,983	95,563	0.5	26,408	48,480	77,574

Table 5 Error indicators for the various networks, meshes and relative residuals

		609fract			*909fract*		
δ	σ	Δ_{head}	Δ_{flux}	Δ_{cons}	Δ_{head}	Δ_{flux}	Δ_{cons}
7	10^{-6}	5.6141e−6	3.3316e−8	6.4713e−4	6.5259e−7	1.0583e−8	1.0849e−3
	10^{-7}	5.6141e−6	3.3316e−8	6.4457e−4	6.5259e−7	1.0583e−8	1.1002e−3
	10^{-8}	5.6141e−6	3.3316e−8	6.4472e−4	6.5259e−7	1.0583e−8	1.1003e−3
2	10^{-6}	3.0565e−6	1.7465e−8	2.9737e−5	4.6689e−7	7.6415e−9	8.9169e−4
	10^{-7}	3.0564e−6	1.7466e−8	2.8968e−5	4.6689e−7	7.6413e−9	9.0033e−4
	10^{-8}	3.0564e−6	1.7466e−8	2.9029e−5	4.6689e−7	7.6414e−9	9.0005e−4
0.5	10^{-6}	1.9031e−6	1.0716e−8	2.6639e−5	2.6142e−7	3.5418e−9	1.8584e−4
	10^{-7}	1.9030e−6	1.0716e−8	2.6829e−5	2.6140e−7	3.5420e−9	1.8786e−4
	10^{-8}	1.9030e−6	1.0716e−8	2.6715e−5	2.6140e−7	3.5420e−9	1.8756e−4
		709fract			*959fract*		
δ	σ	Δ_{head}	Δ_{flux}	Δ_{cons}	Δ_{head}	Δ_{flux}	Δ_{cons}
7	10^{-6}	1.3952e−6	7.0877e−9	5.7534e−3	1.7037e−6	2.1318e−9	2.5212e−3
	10^{-7}	1.3950e−6	7.0862e−9	5.7521e−3	1.7036e−6	2.1296e−9	2.5354e−3
	10^{-8}	1.3950e−6	7.0862e−9	5.7518e−3	1.7036e−6	2.1297e−9	2.5351e−3
2	10^{-6}	1.0801e−6	3.5207e−9	3.4049e−4	1.1070e−6	1.3376e−9	9.6060e−4
	10^{-7}	1.0800e−6	3.5204e−9	3.4102e−4	1.1067e−6	1.3366e−9	9.6625e−4
	10^{-8}	1.0800e−6	3.5204e−9	3.4093e−4	1.1067e−6	1.3365e−9	9.6625e−4
0.5	10^{-6}	6.6866e−7	3.3422e−9	7.1786e−4	6.4611e−7	8.4986e−10	3.5035e−4
	10^{-7}	6.6860e−7	3.3404e−9	7.1523e−4	6.4543e−7	8.4875e−10	3.5232e−4
	10^{-8}	6.6860e−7	3.3403e−9	7.1534e−4	6.4542e−7	8.4871e−10	3.5241e−4

above are reported in Table 5, for all the DFNs and grids and in correspondence of the same values for σ. Finally, Table 6 shows the value of $\bar{\Phi}$ computed for all the network and meshes with the lowest value of the residual and used in the definition of the error indicators. Looking at Table 4 it can be seen that, for all considered meshsizes, the number of iterations required by the method for reaching a prescribed relative residual appears to be, in general, not significantly affected

Table 6 Value of $\bar{\Phi}$ for the considered networks vs meshsize

δ	*609fract*	*709fract*	*909fract*	*959fract*
7	$1.1182e+1$	$1.2672e+1$	3.9928	$4.5195e+1$
2	$1.1263e+1$	$1.2672e+1$	3.9869	$4.4985e+1$
0.5	$1.1286e+1$	$1.2700e+1$	3.9865	$4.4896e+1$

by the presence of critical geometrical features in the network: see for example the case $\sigma = 10^{-6}$, in which the lowest number of iterations is attained on the "difficult" *909fract* network. An exception is provided by the case $\sigma = 10^{-8}$, in which the *909fract* network requires an higher number of iterations. Note however that this is only true on the coarser meshes, and that the value $\sigma = 10^{-8}$ is quite demanding, often corresponding to an over-solving (see later).

The number of iterations appears to be quite independent also of the number of traces and fractures in DFNs, and independent of the meshsize (and thus of the size of the linear system), thanks to the use of preconditioning.

Looking at Table 5 it is possible to notice that the values of the error indicators are very similar in all the considered cases, independently of the geometrical complexities of the networks. Furthermore, it can be noted that pushing the resolution of the linear system to values of the residual lower than 10^{-6} does not provide substantial improvements in the quality of the solution, denoting that the solution rapidly converges to the best solution for a fixed meshsize. A reduction in meshsize, instead, provides, as expected, an improvement in the solution in all the considered cases. Looking at Table 6 we can see that, on the contrary, the value of $\bar{\Phi}$ is quite independent of the mesh.

4.2 Unsteady Flow in DFNs

We end this section with some preliminary data on the resolution of the unsteady advection-diffusion problem in DFNs with the proposed approach. The optimization method is used to compute the Darcy velocity in a given DFN, through the computation of the gradient of the evaluated hydraulic head distribution. At each time-step of the backward Euler method, the optimization method is then used to compute the concentration of a passive scalar quantity C (a pollutant) in the network. Using the same notation for the hydraulic head, at a fixed time frame, let us assume $C_i \in \mathrm{H}^1_{\mathrm{D}}(F_i)$, $\forall i \in \mathscr{I}$, with some prescribed Dirichlet boundary conditions on Γ^D; furthermore, let us assume, for simplicity of notation, homogeneous Neumann boundary conditions on Γ^N and no source terms. Then we have:

$$\min J(C,U) := \sum_{m\in\mathscr{M}} \left(\left\| C_{i|S_m} - C_{j|S_m} \right\|^2_{\mathrm{H}^{\frac{1}{2}}(S_m)} + \left\| U_i^{S_m} + U_j^{S_m} - \alpha \left(C_{i|S_m} + C_{j|S_m} \right) \right\|^2_{\mathrm{H}^{-\frac{1}{2}}(S_m)} \right)$$

such that:

$$\frac{1}{\Delta t}\int_{F_i} C_i v \mathrm{d}\Omega + \int_{F_i} \mu_i \nabla C_i \nabla v \mathrm{d}\Omega + \alpha \sum_{m\in\mathscr{M}} \int_{S_m} C_{i|_{S_m}} v_{|_{S_m}} \mathrm{d}\gamma + \int_{F_i} \beta_i \nabla C_i v \mathrm{d}\Omega =$$

$$\frac{1}{\Delta t}\int_{F_i} C_i^0 v \mathrm{d}\Omega + \sum_{m\in\mathscr{M}} \int_{S_m} U_i^m v_{|_{S_m}} \mathrm{d}\gamma, \quad \forall v \in V_i, \forall i \in \mathscr{I}, \tag{15}$$

where now functions U_i^m act as control variables of the minimization problem for C, μ_i is the diffusivity of the pollutant in F_i, β_i is the velocity field on F_i, Δt is the time-step, and C_i^0 is a term accounting for the solution at the previous time step. Analogously to what described in Sect. 3, the cost functional $J(C, U)$ expresses the mismatch in the continuity and the unbalance of the diffusive flux of the pollutant through the traces of the DFN and it is minimized constrained by the SUPG-stabilized [23] version of (15) on the fractures, expressing the advection-diffusion law for the passive scalar C_i. Results are shown for a 59 fracture DFN, with two different fields of fracture transmissivity: in a first batch of simulations the fracture transmissivity is set to one for all fractures, whereas, in a second set of simulations, the fracture transmissivity is a piecewise constant function with a different value on each fracture. The geometry of the considered network is shown in Fig. 7, left, where the computational mesh and the hydraulic head distribution for the uniform transmissivity field case is also shown, whereas in Fig. 7, right, the transmissivity field in the non homogeneous case is plotted. Figure 8 reports a zoom of the inflow and outflow fractures for the homogeneous transmissivity field, also showing vectors of the computed Darcy velocity. The diffusion of the passive scalar

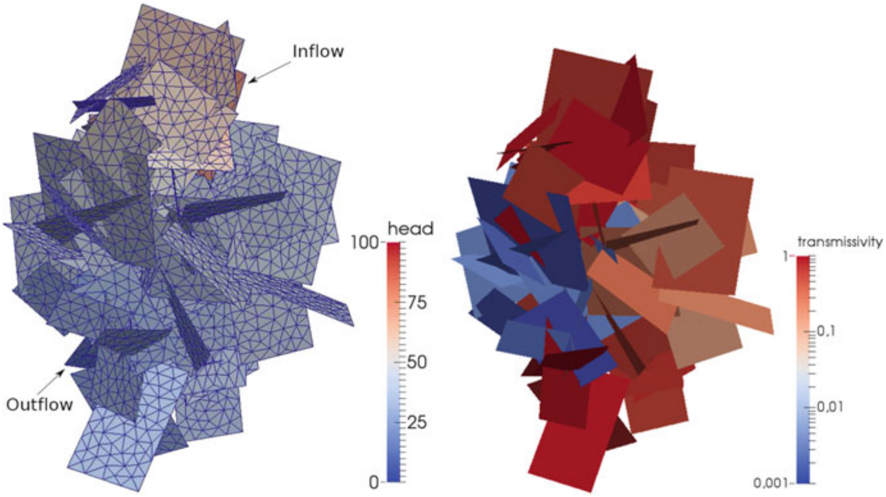


Fig. 7 59 fracture DFN. *Left*: computational mesh and hydraulic head distribution (uniform transmissivity field). *Right*: representation of the non uniform, fracture-constant transmissivity field

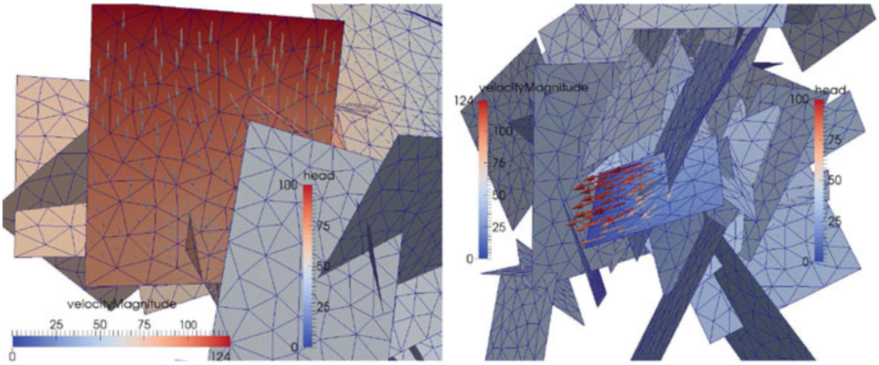


Fig. 8 Velocity field in inflow (*left*) and outflow (*right*) fractures for the 59 fracture DFN. Uniform transmissivity field

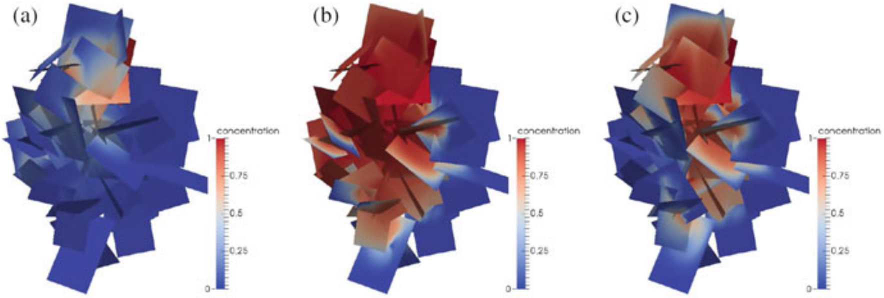


Fig. 9 Pollutant concentration after 10^2 and 10^3 time units for the uniform transmissivity field ((**a**) and (**b**)) and after 10^3 time units for the non-uniform transmissivity field (**c**)

after 10^2 and 10^3 time units for the uniform transmissivity field is shown in Fig. 9a and b: the passive scalar flows from the inflow to the outflow boundary, with an advection-dominated transport, with the diffusivity coefficient for the pollutant set to 10^{-5}. The solution at 10^3 time units for the non uniform transmissivity field is also reported for comparison in Fig. 9c, highlighting the different path followed by the pollutant in the two cases.

Acknowledgements This work has been supported by the Italian MIUR through PRIN research grant 2012HBLYE4_001 “Metodologie innovative nella modellistica differenziale numerica” and by INdAM-GNCS.

References

1. Adler, P.M.: Fractures and Fracture Networks. Kluwer Academic, Dordrecht (1999)
2. Ahmad, B., Alsaedi, A., Brezzi, F., Marini, L.D., Russo, A.: Equivalent projectors for virtual element methods. Comput. Math. Appl. **66**(3), 376–391 (2013). doi:http://dx.doi.org/10.1016/j.camwa.2013.05.015
3. Beirão da Veiga, L., Brezzi, F., Cangiani, A., Manzini, G., Marini, L.D., Russo, A.: Basic principles of virtual element methods. Math. Models Methods Appl. Sci. **23**(1), 199–214 (2013). doi:http://dx.doi.org/10.1142/S0218202512500492
4. Beirão da Veiga, L., Brezzi, F., Marini, L.D.: Virtual elements for linear elasticity problems. SIAM J. Numer. Anal. **51**(2), 794–812 (2013). doi:http://dx.doi.org/10.1137/120874746
5. Beirão da Veiga, L., Brezzi, F., Marini, L.D., Russo, A.: Virtual element method for general second-order elliptic problems on polygonal meshes. Math. Models Methods Appl. Sci. **26**(4), 729–750 (2016). doi:https://doi.org/10.1142/S0218202516500160
6. Beirão da Veiga, L., Lipnikov, K., Manzini, G.: The Mimetic Finite Difference Method for Elliptic Problems. Modeling, Simulation and Applications, vol. 11. Springer, Cham (2014)
7. Belytschko, T., Black, T.: Elastic crack growth in finite elements with minimal remeshing. Int. J. Numer. Methods Eng. **45**, 601–620 (1999)
8. Benedetto, M.F., Berrone, S., Pieraccini, S., Scialò, S.: The virtual element method for discrete fracture network simulations. Comput. Methods Appl. Mech. Eng. **280**(0), 135–156 (2014). doi:http://dx.doi.org/10.1016/j.cma.2014.07.016
9. Benedetto, M.F., Berrone, S., Borio, A.: The virtual element method for underground flow simulations in fractured media. In: Ventura, G., Benvenuti, E. (eds.) Advances in Discretization Methods. SEMA SIMAI Springer Series, vol. 12. Springer International Publishing, Switzerland (2016)
10. Benedetto, M.F., Berrone, S., Scialò, S.: A globally conforming method for solving flow in discrete fracture networks using the virtual element method. Finite Elem. Anal. Des. **109**, 23–36 (2016). doi:http://dx.doi.org/10.1016/j.finel.2015.10.003
11. Berrone, S., Pieraccini, S., Scialò, S.: A PDE-constrained optimization formulation for discrete fracture network flows. SIAM J. Sci. Comput. **35**(2), B487–B510 (2013). doi:http://dx.doi.org/10.1137/120865884
12. Berrone, S., Pieraccini, S., Scialò, S.: On simulations of discrete fracture network flows with an optimization-based extended finite element method. SIAM J. Sci. Comput. **35**(2), A908–A935 (2013). doi:http://dx.doi.org/10.1137/120882883
13. Berrone, S., Pieraccini, S., Scialò, S.: An optimization approach for large scale simulations of discrete fracture network flows. J. Comput. Phys. **256**, 838–853 (2014). doi:http://dx.doi.org/10.1016/j.jcp.2013.09.028
14. Berrone, S., Pieraccini, S., Scialò, S., Vicini, F.: A parallel solver for large scale DFN flow simulations. SIAM J. Sci. Comput. **37**(3), C285–C306 (2015). doi:http://dx.doi.org/10.1137/140984014
15. Berrone, S., Borio, A., Scialò, S.: A posteriori error estimate for a PDE-constrained optimization formulation for the flow in DFNs. SIAM J. Numer. Anal. **54**(1), 242–261 (2016)
16. Berrone, S., Pieraccini, S., Scialò, S.: Towards effective flow simulations in realistic discrete fracture networks. J. Comput. Phys. **310**, 181–201 (2016). doi:http://dx.doi.org/10.1016/j.jcp.2016.01.009
17. Cammarata, G., Fidelibus, C., Cravero, M., Barla, G.: The hydro-mechanically coupled response of rock fractures. Rock Mech. Rock Eng. **40**(1), 41–61 (2007). doi:http://dx.doi.org/10.1007/s00603-006-0081-z
18. Dershowitz, W.S., Fidelibus, C.: Derivation of equivalent pipe networks analogues for three-dimensional discrete fracture networks by the boundary element method. Water Resour. Res. **35**, 2685–2691 (1999). doi:http://dx.doi.org/10.1029/1999WR900118
19. de Dreuzy, J.R., Pichot, G., Poirriez, B., Erhel, J.: Synthetic benchmark for modeling flow in 3D fractured media. Comput. Geosci. **50**(0), 59–71 (2013)

20. Fidelibus, C., Cammarata, G., Cravero, M.: Hydraulic characterization of fractured rocks. In: Abbie, M., Bedford, J.S. (eds.) Rock Mechanics: New Research. Nova Science Publishers Inc., New York (2009)
21. Formaggia, L., Fumagalli, A., Scotti, A., Ruffo, P.: A reduced model for Darcy's problem in networks of fractures. ESAIM: Math. Model. Numer. Anal. **48**, 1089–1116 (2014). doi:http://dx.doi.org/0.1051/m2an/2013132
22. Formaggia, L., Antonietti, P., Panfili, P., Scotti, A., Turconi, L., Verani, M., Cominelli, A.: Optimal techniques to simulate flow in fractured reservoir. In: ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery (2014)
23. Franca, L.P., Frey, S.L., Hughes, T.J.R.: Stabilized finite element methods: I. Application to the advective-diffusive model. Comput. Methods Appl. Mech. Eng. **95**, 253–276 (1992)
24. Fries, T.P.: A corrected XFEM approximation without problems in blending elements. Int. J. Numer. Methods Eng. **75**, 503–532 (2008). doi:http://dx.doi.org/10.1002/nme.2259
25. Fries, T.P., Belytschko, T.: The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Methods Eng. **84**(3), 253–304 (2010). doi:http://dx.doi.org/10.1002/nme.2914
26. Fumagalli, A., Scotti, A.: A numerical method for two-phase flow in fractured porous media with non-matching grids. Adv. Water Resour. **62**, 454–464 (2013). doi:http://dx.doi.org/10.1016/j.advwatres.2013.04.001
27. Fumagalli, A., Scotti, A.: An efficient XFEM approximation of Darcy flows in arbitrarily fractured porous media. Oil Gas Sci. Technol. - Rev. d'IFP Energies Nouv. **69**(4), 555–564 (2014). doi:http://dx.doi.org/10.2516/ogst/2013192
28. Hyman, J., Gable, C., Painter, S., Makedonska, N.: Conforming Delaunay triangulation of stochastically generated three dimensional discrete fracture networks: a feature rejection algorithm for meshing strategy. SIAM J. Sci. Comput. **36**, A1871–A1894 (2014). doi:http://dx.doi.org/10.1137/130942541
29. Jaffré, J., Roberts, J.E.: Modeling flow in porous media with fractures; discrete fracture models with matrix-fracture exchange. Numer. Anal. Appl. **5**(2), 162–167 (2012)
30. Karimi-Fard, M., Durlofsky, L.J.: Unstructured adaptive mesh refinement for flow in heterogeneous porous media. In: ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery (2014)
31. Nœtinger, B.: A quasi steady state method for solving transient Darcy flow in complex 3D fractured networks accounting for matrix to fracture flow. J. Comput. Phys. **283**, 205–223 (2015). doi:http://dx.doi.org/10.1016/j.jcp.2014.11.038
32. Nœtinger, B., Jarrige, N.: A quasi steady state method for solving transient Darcy flow in complex 3D fractured networks. J. Comput. Phys. **231**(1), 23–38 (2012). doi:http://dx.doi.org/10.1016/j.jcp.2011.08.015
33. Osher, S., Fedkiw, R.: Level Set Methods and Dynamic Implicit Surfaces. Applied Mathematical Sciences, vol. 153. Springer, New York (2003)
34. Pichot, G., Erhel, J., de Dreuzy, J.: A mixed hybrid mortar method for solving flow in discrete fracture networks. Appl. Anal. **89**, 1629–643 (2010). doi:http://dx.doi.org/10.1080/00036811.2010.495333
35. Pichot, G., Erhel, J., de Dreuzy, J.: A generalized mixed hybrid mortar method for solving flow in stochastic discrete fracture networks. SIAM J. Sci. Comput. **34**, B86–B105 (2012). doi:http://dx.doi.org/10.1137/100804383
36. Pichot, G., Poirriez, B., Erhel, J., de Dreuzy, J.R.: A Mortar BDD method for solving flow in stochastic discrete fracture networks. In: Domain Decomposition Methods in Science and Engineering XXI. Lecture Notes in Computational Science and Engineering, pp. 99–112. Springer, Cham (2014)
37. Vohralík, M., Maryška, J., Severýn, O.: Mixed and nonconforming finite element methods on a system of polygons. Appl. Numer. Math. **51**, 176–193 (2007)

A Review of the XFEM-Based Approximation of Flow in Fractured Porous Media

Bernd Flemisch, Alessio Fumagalli, and Anna Scotti

Abstract This paper presents a review of the available mathematical models and corresponding non-conforming numerical approximations which describe single-phase fluid flow in a fractured porous medium. One focus is on the geometrical difficulties that may arise in realistic simulations such as intersecting and immersed fractures. Another important aspect is the choice of the approximation spaces for the discrete problem: in mixed formulations, both the Darcy velocity and the pressure are considered as unknowns, while in classical primal formulations, a richer space for the pressure is considered and the Darcy velocity is computed a posteriori. In both cases, the extended finite element method is used, which allows for a complete geometrical decoupling among the fractures and rock matrix grids. The fracture geometries can thus be independent of the underlying grid thanks to suitable enrichments of the spaces that are able to represent possible jumps of the solution across the fractures. Finally, due to the dimensional reduction, a better approximation of the resulting boundary conditions for the fractures is addressed.

1 Introduction

The simulation of subsurface flow is of great importance for a large number of applications ranging from the production of energy (oil and gas reservoirs, geothermal energy) to the management of water resources, or the safe storage of atomic waste and carbon dioxide. Fractures are present in porous media at a variety of scales. Large fractures and faults in particular are very relevant for the flow since they can either act as barriers (in the case of impermeable faults) or preferential pathways for the flow (in the case of permeable fractures). Opposed to small-scale

---

B. Flemisch
Department of Hydromechanics and Modelling of Hydrosystems, University of Stuttgart,
Pfaffenwaldring 61, 70569 Stuttgart, Germany
e-mail: bernd@iws.uni-stuttgart.de

A. Fumagalli • A. Scotti (✉)
MOX, Dipartimento di Matematica, Politecnico di Milano, via Bonardi 9, 20133 Milan, Italy
e-mail: alessio.fumagalli@polimi.it; anna.scotti@polimi.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_3

fractures, that can be accounted for by upscaling of the permeability, large features should be explicitly included in the model to reproduce their non-local effects on the flow. Thanks to the developments of numerical methods and computing power direct numerical simulations of fracture networks are replacing or complementing multi-continua approaches such as dual-porosity/permeability.

Fractures and faults are three-dimensional regions characterized by a different porosity and permeability with respect to the surrounding porous matrix. However, thanks to their small aperture compared to the typical length and the size of the domain, they are usually represented as $(N-1)$-dimensional interfaces immersed in a N-dimensional matrix. From a computational viewpoint, this avoids the need for an extremely fine grid to resolve the width of fractures, that are now replaced by discontinuity surfaces where a suitable reduced $(N-1)$-dimensional problem is solved and coupled with the surrounding flow. However, the complexity of geological structures remains one of the main challenges in large-scale numerical simulations. Indeed, the data for the construction of the model are usually given as a large number of possibly intersecting surfaces, called horizons, that separate layers with different mechanic and hydraulic properties, and a set of surfaces that represent faults and fractures. In this framework, the construction of a grid which is conforming with all the aforementioned features is a difficult task, [20], whose outcome could be a grid that is either too refined to be used, or with low quality elements. Since in these realistic cases the construction of a high-quality grid that honours the geometry of hundreds or thousands of fractures is a challenging task, two alternative approaches are possible:

- to develop numerical methods that are accurate and robust even for very distorted grids, such as the Mimetic Finite Difference Methods [8];
- to allow the fractures to cross a fairly regular and coarse grid in arbitrary ways, and to employ the eXtended Finite Element Method (XFEM) to account for the solution discontinuities within elements.

This paper presents the second approach, reviewing the recent literature on the application of the XFEM to the simulation of flow in fractured porous media, focusing on single-phase flow in the presence of one or more, possibly intersecting, fractures.

The XFEM has been successfully used for the simulation of crack mechanics for a long time, [22, 23, 44], while its application to flow in fractured media is a recent development. Allowing for non-matching grids with respect to the fracture network can be advantageous in geological problems since not only it avoids the burden of computing a conforming grid, but it avoids the need for re-meshing in the case of uncertain geometry, i.e. one could perform simulations of different scenarios with different fracture configurations with the same background grid.

The enrichment of the finite element spaces should be able to effectively represent discontinuities in the pressure and in the flux across fractures: pressure jumps arise in the case of impermeable interfaces, while a discontinuous flux can be observed due to the fact that fluid can enter the fractures and flow along them.

The development and the analysis of $(N-1)$-dimensional models for fractures for single-phase flow have been extensively addressed in [3, 7, 13, 29, 42], where the fracture flow equations and the proper interface conditions across the fracture have been first derived, and the continuous and discrete problems have been studied in their mixed formulation. However, in the aforementioned works, the computational grid of the porous domain is considered to be matching with the fracture, i.e. the fracture is the (conforming) interface between two mesh blocks, possibly with different resolution. Similarly, in more recent works, this type of space discretization has been employed to describe the flow in faulted sedimentary basins, coupled with a double-layer model for the fault, see [26, 52]. However, while non-conforming meshes on the interface could be dealt with by mortaring, this does not allow the fractures to cut the elements of the grid.

The use of XFEM to deal with fractures as non-matching, immersed interfaces, was first introduced in 2011, for the single phase case, in [19, 41]. In the former, the concept of EFEM (Enriched Finite Element Method) is applied to the primal formulation with suitable enrichments for the pressure, while in the latter the mixed formulation of the problem is considered, and the authors employ concepts borrowed from [39] to enrich both the pressure and the Darcy velocity spaces. In the same years, XFEM have been applied for the discretization of the primal formulation in [50].

The application to the case of several intersecting fractures is discussed in [11, 27] in the case of an impervious surrounding medium, with slightly different coupling conditions at the intersections based on different assumptions on the fractures permeability. The coupled problem, in the mixed and primal formulation respectively, are considered in [30, 31, 51]. As concerns the physics of flow in porous media, some works consider also the case of passive transport of solutes in fractured porous media, [32], and two-phase flow [34] by means of the XFEM. Moreover, the coupling of Darcy flow with fracture mechanics (opening, propagation) is addressed in [37, 48].

Regarding the analysis of the method, the inf-sup stability of XFEM applied to Darcy flow in porous media has been proven, under suitable conditions, in [19, 21].

This paper is organized as follows. In Sect. 2, we introduce the mathematical model for single-phase flow in fractured porous media in the equi-dimensional case, and derive the corresponding hybrid-dimensional, or reduced, model. Intersecting fractures are considered, as well as the task of assigning boundary conditions for the fractures. In Sect. 3, we present the numerical discretization techniques for the problem in primal and dual mixed form, with a focus on the ad-hoc enrichments at intersections and tips, and on the approximation of coupling terms. Section 4 is dedicated to solvers for the resulting linear system, in particular to conditioning issues and the choice of iterative vs. monolithic approaches. Finally, Sect. 5 is devoted to some concluding remarks and future perspectives.

2 Governing Equations

In this section, we present the mathematical model of single-phase flow in porous media, focusing our attention on the description of the fractures. We start considering the standard Darcy law and mass balance in an equi-dimensional setting, where the fractures are N-dimensional regions embedded in an N-dimensional porous matrix. Then we introduce the so called reduced, or hybrid-dimensional, models to handle fractures as objects of effective lower dimension, intersections and branching of fractures, and proper boundary conditions to prescribe at the fracture tips and on the cut matrix boundaries.

2.1 *Equi-Dimensional Models*

We consider an inert and at rest porous medium which can be modeled as a bounded, connected, and open set $\mathcal{D} \subset \mathbb{R}^N$, $N{=}2$ or 3. We assume that the medium is saturated with a single incompressible fluid phase that is composed of a unique component, e.g., water. The boundary, which is required to be regular enough for the forthcoming assumptions, is indicated by $\partial\mathcal{D}$ with outward unit normal $\boldsymbol{n}_{\mathcal{D}}$. Let us assume that $\mathcal{D}$ contains several fractures, that all together constitute a single domain Γ of spatial dimension N such that $\Gamma \subset \mathcal{D}$, which is a possibly unconnected, open subset of $\mathcal{D}$. The fracture network can also be seen as the union of fracture branches Γ_i. The surrounding porous rock, namely, the remaining part of $\mathcal{D}$, is called $\Omega := \mathcal{D} \setminus \overline{\Gamma}$. The outer boundary of the rock matrix is indicated by $\partial\Omega = \overline{\Omega} \cap \partial\mathcal{D}$, while the outer boundary of the fracture network is indicated by $\partial\Gamma$ and defined by $\partial\Gamma := \overline{\Gamma} \cap \partial\mathcal{D}$. Moreover, the internal part of the fracture boundary, namely, the interface between the fracture domain and the surrounding medium, is indicated by γ and defined as $\gamma := \overline{\Gamma} \cap \overline{\Omega}$. For each fracture branch, we call its own part of this internal boundary γ_i. The unit normal, pointing out of Γ into Ω, is indicated by $\boldsymbol{n}$. We suppose that for each fracture branch there exists a central axis $\hat{\gamma}_i$, which is a non self-intersecting $(N{-}1)$-dimensional surface, such that a fracture branch can be described as

$$\Gamma_i = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \boldsymbol{x} = \boldsymbol{s} + r\boldsymbol{n}_i,\ \boldsymbol{s} \in \hat{\gamma}_i,\ |r| < \frac{d_i}{2} \right\}, \tag{1}$$

where d_i is the aperture of Γ_i, which may depend on the curvilinear abscissa $\boldsymbol{s}$, and $\boldsymbol{n}_i$ is the unit normal associated with the central axis $\hat{\gamma}_i$. We assume that the apertures d_i are small compared to other characteristic dimensions of the fractures. With definition (1) the fracture domain is composed by $\overline{\Gamma} = \cup\overline{\Gamma_i}$, but note that Γ_i may intersect each other with a non-null intersection. See Fig. 1 for an illustration of the aforementioned notation.

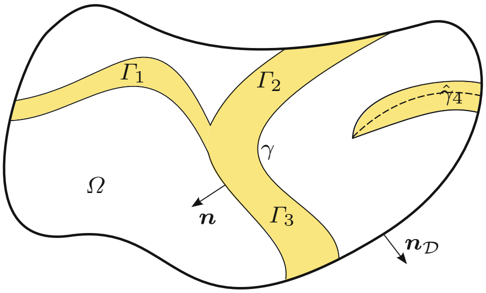


Fig. 1 Notation for a general configuration of fractures in the equi-dimensional model

In this work we assume that the fractures are filled by a porous medium themselves such that Darcy flow takes place in both the rock matrix and fractures. In the relevant case of open fractures the lubrication model could be used, see for instance [53].

2.1.1 Dual Formulation

In this part, our objective is to compute the steady-state pressure field p and the Darcy velocity field, or macroscopic velocity, $\boldsymbol{u}$ in the entire porous domain $\mathcal{D}$. To this purpose, following for example [9], we employ the law of mass conservation together with Darcy's law and, to ease the notation, we assume homogeneous boundary conditions for the pressure on the whole boundary. The system of equations for the porous matrix Ω is given by

$$\begin{cases} \nabla \cdot \boldsymbol{u} = f \\ \boldsymbol{u} + \Lambda \nabla p = \boldsymbol{0} \end{cases} \text{in } \Omega, \quad \begin{cases} p = 0 \end{cases} \text{on } \partial\Omega, \tag{2a}$$

where the scalar source term f represents a possible volume source or sink and Λ denotes the symmetric and positive definite permeability tensor in Ω. To simplify the notation we consider a permeability tensor that is already scaled with the viscosity. Coupled with (2a), a similar system of equations can be considered for the fracture network. The data and unknowns related to the fractures are indicated with a subscript f. We obtain

$$\begin{cases} \nabla \cdot \boldsymbol{u}_{\mathrm{f}} = f_{\mathrm{f}} \\ \boldsymbol{u}_{\mathrm{f}} + \Lambda_{\mathrm{f}} \nabla p_{\mathrm{f}} = \boldsymbol{0} \end{cases} \text{in } \Gamma, \quad \begin{cases} p_{\mathrm{f}} = 0 \end{cases} \text{on } \partial\Gamma. \tag{2b}$$

Following [42] we require that the permeability tensor in the fracture system, for each fracture branch, can be written as $\Lambda_{\mathrm{f},i} = \lambda_{i,\boldsymbol{n}} \boldsymbol{N}_i + \lambda_{i,\boldsymbol{\tau}} \boldsymbol{T}_i$, where the projection matrix $\boldsymbol{N}_i$ in the direction normal to $\hat{\gamma}_i$ and the projection matrix $\boldsymbol{T}_i$ in the direction tangential to $\hat{\gamma}_i$ are defined as follows:

$$\boldsymbol{N}_i := \boldsymbol{n}_i \otimes \boldsymbol{n}_i \quad \text{and} \quad \boldsymbol{T}_i := \boldsymbol{I} - \boldsymbol{N}_i.$$

To couple the systems (2a) and (2b) we consider the following classical interface conditions, namely

$$\begin{cases} p = p_\mathrm{f} \\ \boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{u}_\mathrm{f} \cdot \boldsymbol{n} \end{cases} \quad \text{on } \gamma. \tag{2c}$$

Combining (2a)–(2c), we obtain the strong problem formulation in its dual form.

Problem 1 (Dual Equi-Dimensional Strong Formulation) Find velocity fields $\boldsymbol{u}, \boldsymbol{u}_\mathrm{f}$ and pressure fields p, p_f such that (2) is fulfilled.

The proof of the well-posedness of Problem 1 in its mixed weak form can be found in a number of references, such as [15, 24, 46, 49].

2.1.2 Primal Formulation

A common formulation for single-phase porous-media flow is the so-called primal formulation, that can be obtained inserting Darcy's law for matrix and fracture domain, namely, the second lines of (2a), (2b) into the mass-balance equations, namely, the first lines of (2a), (2b), as well as into the flux-coupling condition, namely, the second line of (2c). In particular, for the matrix domain Ω, we obtain

$$\begin{cases} -\nabla \cdot (\Lambda \nabla p) = f & \text{in } \Omega, \\ p = 0 & \text{on } \partial\Omega, \end{cases} \tag{3a}$$

while for the fracture domain Γ, we have

$$\begin{cases} -\nabla \cdot (\Lambda_\mathrm{f} \nabla p_\mathrm{f}) = f_\mathrm{f} & \text{in } \Gamma, \\ p_\mathrm{f} = 0 & \text{on } \partial\Gamma, \end{cases} \tag{3b}$$

coupled by

$$\begin{cases} p = p_\mathrm{f} \\ \Lambda \nabla p \cdot \boldsymbol{n} = \Lambda_\mathrm{f} \nabla p_\mathrm{f} \cdot \boldsymbol{n} \end{cases} \quad \text{on } \gamma. \tag{3c}$$

The problem can, in this case, be cast as follows.

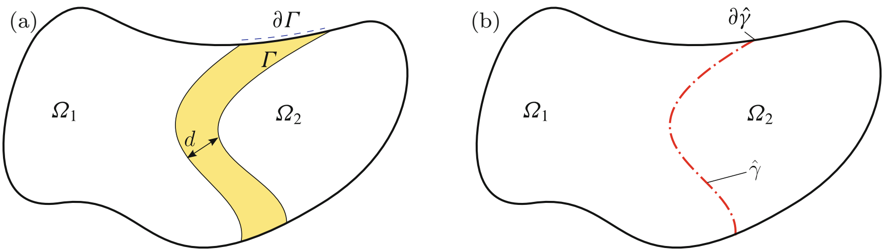


Fig. 2 Model domains with fracture. (**a**) Equi-dimensional model domain. (**b**) Hybrid-dimensional model domain which includes a lower dimensional fracture

Problem 2 (Primal Equi-Dimensional Strong Formulation) Find pressure fields p, p_f such that (3) is fulfilled.

The well-posedness of Problem 2 and its weak form can be found in any standard textbook on partial differential equations or finite elements (Fig. 2).

2.2 Hybrid-Dimensional Models

In this section, we present the hybrid-dimensional model, or reduced model, in the case of single fracture dividing the domain in two unconnected parts. We refer to [2, 6, 19, 25, 42] for a detailed presentation. The derivation of the model is based on its dual formulation, however, we present also its primal formulation, obtained with a "post-processing". For both formulations, we briefly introduce their weak formulation to be used for the numerical discretization.

During the process, we substitute the fracture Γ by its centre line $\hat{\gamma}$ and the surrounding porous medium is enlarged to fill the gap. In practical cases, this step is very seldom performed since the fracture geometry is directly given as an object of codimension one. The Darcy equations (2a) for the rock matrix are thus the same as in the equi-dimensional case and we focus only on the Eqs. (2b) for the fracture. Since Ω is split in two parts $\Omega_{1,2}$, we define $\boldsymbol{n} = \boldsymbol{n}_1 = -\boldsymbol{n}_2$, where $\boldsymbol{n}_i$ is the outward-pointing normal of Ω_i. Let us introduce the normal and tangential divergence and gradient on the fracture: given two regular functions $\boldsymbol{a}$ and a, respectively vector- and scalar-valued, we define

$$\nabla \cdot \boldsymbol{a} = \nabla_{\boldsymbol{n}} \cdot \boldsymbol{a} + \nabla_{\boldsymbol{\tau}} \cdot \boldsymbol{a} \quad \text{with} \quad \nabla_{\boldsymbol{n}} \cdot \boldsymbol{a} := \boldsymbol{N} : \nabla \boldsymbol{a} \quad \text{and} \quad \nabla_{\boldsymbol{\tau}} \cdot \boldsymbol{a} := \boldsymbol{T} : \nabla \boldsymbol{a},$$

$$\nabla a = \nabla_{\boldsymbol{n}} a + \nabla_{\boldsymbol{\tau}} a \quad \text{with} \quad \nabla_{\boldsymbol{n}} a := \boldsymbol{N} \nabla a \quad \text{and} \quad \nabla_{\boldsymbol{\tau}} a := \boldsymbol{T} \nabla a.$$

2.2.1 Dual Formulation

We decompose the Darcy velocity in the fracture into its normal part $\boldsymbol{u}_{\mathrm{f},\boldsymbol{n}} := \boldsymbol{N}\boldsymbol{u}_{\mathrm{f}}$ and tangential part $\boldsymbol{u}_{\mathrm{f},\boldsymbol{\tau}} := \boldsymbol{T}\boldsymbol{u}_{\mathrm{f}}$, such that $\boldsymbol{u}_{\mathrm{f}} = \boldsymbol{u}_{\mathrm{f},\boldsymbol{n}} + \boldsymbol{u}_{\mathrm{f},\boldsymbol{\tau}}$. We consider first the conservation equation which is integrated along the normal direction of $\hat{\gamma}$ for the fracture aperture d. We obtain a conservation equation in the tangential space of $\hat{\gamma}$ for the reduced flux $\hat{\boldsymbol{u}} := \int_{-d/2}^{d/2} \boldsymbol{u}_{\mathrm{f},\boldsymbol{\tau}}$ which involves also the contribution of the incoming flux from the surrounding porous medium, namely,

$$\nabla_{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{u}} = \hat{f} + [\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} \quad \text{in } \hat{\gamma}, \tag{4}$$

where the reduced source term is defined as $\hat{f} := \int_{-d/2}^{d/2} f_{\mathrm{f}}$. In (4) we made use of the jump operator defined as $[\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} := \boldsymbol{u}_1 \cdot \boldsymbol{n} - \boldsymbol{u}_2 \cdot \boldsymbol{n}$, with an abuse of notation for the normal $\boldsymbol{n}$. We consider now the Darcy equation projected on the tangential space of $\hat{\gamma}$ and integrated in normal direction for the aperture of the fracture, obtaining

$$\hat{\boldsymbol{u}} + \hat{\lambda} \nabla_{\boldsymbol{\tau}} \hat{p} = \boldsymbol{0} \quad \text{in } \hat{\gamma}, \tag{5}$$

where $\hat{p}$ is the reduced pressure in the fracture, defined as $\hat{p} := \frac{1}{d} \int_{-d/2}^{d/2} p_{\mathrm{f}}$. In the previous equation, $\hat{\lambda}$ is the effective permeability in tangential direction, defined as $\hat{\lambda} := d\lambda_{\mathrm{f},\boldsymbol{\tau}}$. Finally, projecting Darcy's law on the normal space of the fracture and integrating in normal direction on the first and on the second half of the aperture, we end up with coupling conditions between the lower-dimensional fracture and the rock matrix. Using a suitable approximation of the integral of $\boldsymbol{u}_i \cdot \boldsymbol{n}$, as discussed in [7, 42], we get

$$\begin{cases} \xi \boldsymbol{u}_1 \cdot \boldsymbol{n} + (1-\xi) \boldsymbol{u}_2 \cdot \boldsymbol{n} = 2\lambda_{\hat{\gamma}} (p_1 - \hat{p}) \\ \xi \boldsymbol{u}_2 \cdot \boldsymbol{n} + (1-\xi) \boldsymbol{u}_1 \cdot \boldsymbol{n} = 2\lambda_{\hat{\gamma}} (\hat{p} - p_2) \end{cases} \quad \text{on } \hat{\gamma}, \tag{6}$$

where $\lambda_{\hat{\gamma}}$ is the effective permeability in normal direction of the fracture, defined as $\lambda_{\hat{\gamma}} := \lambda_{\mathrm{f},\boldsymbol{n}}/d$. Moreover, $\xi \in (0.5, 1]$ is a closure parameter related to the pressure cross profile in the fracture, see the aforementioned works for more details. Considering (2a) coupled with (4)–(6) we obtain the following hybrid-dimensional problem.

Problem 3 (Dual Hybrid-Dimensional Strong Formulation) Find $(\boldsymbol{u}_i, p_i)$ for $i = 1, 2$ and $(\hat{\boldsymbol{u}}, \hat{p})$ such that

$$\begin{cases} \nabla \cdot \boldsymbol{u}_i = f_i & \text{in } \Omega_i, \\ \boldsymbol{u}_i + \Lambda_i \nabla p_i = \boldsymbol{0} & \\ p_i = 0 & \text{on } \partial\Omega_i, \end{cases} \quad \text{and} \quad \begin{cases} \nabla_{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{u}} = \hat{f} + [\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} & \text{in } \hat{\gamma}, \\ \hat{\boldsymbol{u}} + \hat{\lambda} \nabla_{\boldsymbol{\tau}} \hat{p} = \boldsymbol{0} & \\ \hat{p} = 0 & \text{on } \partial\hat{\gamma}, \end{cases} \tag{7a}$$

with interface conditions

$$\begin{cases} \xi \boldsymbol{u}_1 \cdot \boldsymbol{n} + (1-\xi)\boldsymbol{u}_2 \cdot \boldsymbol{n} = 2\lambda_{\hat{\gamma}} \left(p_1 - \hat{p}\right) \\ \xi \boldsymbol{u}_2 \cdot \boldsymbol{n} + (1-\xi)\boldsymbol{u}_1 \cdot \boldsymbol{n} = 2\lambda_{\hat{\gamma}} \left(\hat{p} - p_2\right) \end{cases} \quad \text{on } \hat{\gamma}. \tag{7b}$$

An alternative form of the interface conditions (7b), introduced in [18], is

$$\begin{cases} \{\!\{\boldsymbol{u} \cdot \boldsymbol{n}\}\!\}_{\hat{\gamma}} = \lambda_{\hat{\gamma}} \, [\![p]\!]_{\hat{\gamma}} \\ \xi_0 \, [\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} = \lambda_{\hat{\gamma}} \left(\{\!\{p\}\!\}_{\hat{\gamma}} - \hat{p}\right) \end{cases} \quad \text{on } \hat{\gamma}, \tag{7b-bis}$$

with $\xi_0 = 4/(2\xi - 1)$ and where we have used the average operators $\{\!\{p\}\!\}_{\hat{\gamma}} := \frac{1}{2}\left(p_1 + p_2\right)$ and $\{\!\{\boldsymbol{u} \cdot \boldsymbol{n}\}\!\}_{\hat{\gamma}} := \frac{1}{2}\left(\boldsymbol{u}_1 \cdot \boldsymbol{n} + \boldsymbol{u}_2 \cdot \boldsymbol{n}\right)$, as well as the jump operator for the pressure $[\![p]\!]_{\hat{\gamma}} := p_1 - p_2$.

We now introduce the weak formulation of the reduced problem, which will be useful to present the XFEM in Sect. 3. For a detailed presentation of the suitable functional spaces refer to [7]. We start by introducing the following bilinear forms and functionals for the rock matrix

$$a_{\mathrm{d}}(\boldsymbol{u}, \boldsymbol{v}) := \sum_i \left(H\boldsymbol{u}_i, \boldsymbol{v}_i\right)_{\Omega_i} + \left(\eta_{\hat{\gamma}} \{\!\{\boldsymbol{u} \cdot \boldsymbol{n}\}\!\}_{\hat{\gamma}}, \{\!\{\boldsymbol{v} \cdot \boldsymbol{n}\}\!\}_{\hat{\gamma}}\right)_{\hat{\gamma}}$$
$$+\xi_0 \left(\eta_{\hat{\gamma}} \, [\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}}, [\![\boldsymbol{v} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}}\right)_{\hat{\gamma}}, \ \text{with}\, i = 1, 2,$$

with $H := \Lambda^{-1}$, $\eta_{\hat{\gamma}} := \lambda_{\hat{\gamma}}^{-1}$ the inverse of the permeabilities. The bilinear form and the functional which include the source term and possibly boundary conditions read

$$b_{\mathrm{d}}(p, \boldsymbol{v}) := -\sum_i \left(p_i, \nabla \cdot \boldsymbol{v}_i\right)_{\Omega_i} \quad \text{and} \quad F(q) := \sum_i \left(f_i, q_i\right)_{\Omega_i}.$$

The weak formulation for the fracture requires to introduce the following bilinear forms and functional

$$\hat{a}_{\mathrm{d}}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}) := \left(\hat{\eta}\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}\right)_{\hat{\gamma}}, \quad \hat{b}_{\mathrm{d}}(\hat{p}, \hat{\boldsymbol{v}}) := -\left(\hat{p}, \nabla_{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{v}}\right)_{\hat{\gamma}} \quad \text{and} \quad \hat{F}(\hat{q}) := \left(\hat{f}, \hat{q}\right)_{\hat{\gamma}}$$

with $\hat{\eta} := \hat{\lambda}^{-1}$ the inverse of the effective tangential permeability. Finally the bilinear form which couples the fracture and the surrounding porous medium

$$c_{\mathrm{d}}(\boldsymbol{u}, \hat{q}) := \left([\![\boldsymbol{u} \cdot \boldsymbol{n}]\!]_{\hat{\gamma}}, \hat{q}\right)_{\hat{\gamma}}.$$

The weak formulation of (7) is given as follows.

Problem 4 (Dual Hybrid-Dimensional Weak Formulation) Find $(\boldsymbol{u}_i, p_i)$ for $i = 1, 2$ and $(\hat{\boldsymbol{u}}, \hat{p})$ respecting the given boundary conditions such that

$$\begin{cases} a_{\mathrm{d}}(\boldsymbol{u}, \boldsymbol{v}) + b_{\mathrm{d}}(p, \boldsymbol{v}) + c_{\mathrm{d}}(\hat{p}, \boldsymbol{v}) = 0 \\ b_{\mathrm{d}}(q, \boldsymbol{u}) = F(q) \end{cases} \quad \text{and} \quad \begin{cases} \hat{a}_{\mathrm{d}}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}) + \hat{b}_{\mathrm{d}}(\hat{p}, \hat{\boldsymbol{v}}) = 0 \\ \hat{b}_{\mathrm{d}}(\hat{q}, \hat{\boldsymbol{u}}) - c_{\mathrm{d}}(\hat{q}, \boldsymbol{u}) = \hat{F}(\hat{q}) \end{cases},$$

for all test functions $\boldsymbol{v}$, q, $\hat{\boldsymbol{v}}$ and $\hat{q}$ defined in their proper spaces.

2.2.2 Primal Formulation

As for the equi-dimensional setting, a primal formulation can be derived by inserting Darcy's laws into the mass balance equations in (7a).

Problem 5 (Primal Hybrid-Dimensional Strong Formulation) Find p_i for $i = 1, 2$ and $\hat{p}$ such that

$$\begin{cases} -\nabla \cdot \Lambda_i \nabla p_i = f_i & \text{in } \Omega_i, \\ p_i = 0 & \text{on } \partial\Omega_i, \end{cases} \quad \text{and} \quad \begin{cases} -\nabla_{\boldsymbol{\tau}} \cdot \hat{\lambda} \nabla_{\boldsymbol{\tau}} \hat{p} = \hat{f} - [\![\Lambda \nabla p \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} & \text{in } \hat{\gamma}, \\ \hat{p} = 0 & \text{on } \partial\hat{\gamma}, \end{cases} \tag{8a}$$

with interface conditions (7b-bis) reformulated as

$$\begin{cases} -\{\!\{\Lambda \nabla p \cdot \boldsymbol{n}\}\!\}_{\hat{\gamma}} = \lambda_{\hat{\gamma}} [\![p]\!]_{\hat{\gamma}} \\ -\xi_0 [\![\Lambda \nabla p \cdot \boldsymbol{n}]\!]_{\hat{\gamma}} = \lambda_{\hat{\gamma}} \left(\{\!\{p\}\!\}_{\hat{\gamma}} - \hat{p}\right) \end{cases} \quad \text{on } \hat{\gamma}. \tag{8b}$$

Proceeding to the weak formulation of the primal problem, we define the matrix bilinear form

$$a_{\mathrm{p}}(p, q) = \sum_i (\Lambda_i \nabla p_i, \nabla q_i)_{\Omega_i} + \left(\xi_0 \lambda_{\hat{\gamma}} \{\!\{p\}\!\}_{\hat{\gamma}}, \{\!\{q\}\!\}_{\hat{\gamma}}\right)_{\hat{\gamma}} + \left(\lambda_{\hat{\gamma}} [\![p]\!]_{\hat{\gamma}}, [\![q]\!]_{\hat{\gamma}}\right)_{\hat{\gamma}}$$

and the fracture bilinear form

$$\hat{a}_{\mathrm{p}}(\hat{p}, \hat{q}) = \left(\hat{\lambda} \nabla_{\boldsymbol{\tau}} \hat{p}, \nabla_{\boldsymbol{\tau}} \hat{q}\right)_{\hat{\gamma}} + \left(\xi_0 \lambda_{\hat{\gamma}} \hat{p}, \hat{q}\right)_{\hat{\gamma}}.$$

The coupling between matrix and fracture is accounted for by the bilinear form

$$c_{\mathrm{p}}(p, \hat{q}) = \left(\xi_0 \lambda_{\hat{\gamma}} \{\!\{p\}\!\}_{\hat{\gamma}}, \hat{q}\right)_{\hat{\gamma}}.$$

This allows to obtain the weak formulation of Problem 5.

Problem 6 (Primal Hybrid-Dimensional Weak Formulation) Find p_i for $i = 1, 2$ and $\hat{p}$ respecting the given boundary conditions such that

$$\begin{cases} a_{\mathrm{p}}(p, q) - c_{\mathrm{p}}(q, \hat{p}) = F(q) \\ -c_{\mathrm{p}}(p, \hat{q}) + \hat{a}_{\mathrm{p}}(\hat{p}, \hat{q}) = \hat{F}(\hat{q}) \end{cases},$$

for all test functions q and $\hat{q}$ defined in their proper spaces.

The analysis of Problem 6 is straightforward and presented in, for example, [50]. However, the situation becomes more involved if the fracture is allowed to end inside the interior of the matrix, yielding a non-Lipschitz domain Ω. A rigorous mathematical analysis of this setup is carried out in [37].

2.3 Branching and Intersections

In this part, we present several strategies to model the intersection of fractures. This is an important aspect, since the complex nature of networks of possibly heterogeneous fractures requires an appropriate treatment to avoid un-physical results. In a crossing, however, the different properties of every fracture branch can overlap and a unique association of properties is not always possible, so that, in general, new properties have to be defined for the crossing area, based on physical arguments to be provided by the modeler. If there is a crossing of fractures with very different permeabilities, one fracture always dominates a crossing from a geological point of view. For example, if there exists a highly permeable fracture which becomes intersected over time by an almost impermeable fracture, the crossing permeability is more likely to be almost impermeable than highly conductive or averaged. It is then neither a realistic choice to always average the permeabilities in a crossing nor to neglect the connection between different fractures. In the forthcoming reduced models, to simplify the notation, we focus our attention on a single fractures intersection inside the porous domain where several fracture branches $\hat{\gamma}_k$ meet. In all the subsequent cases, the model for the flow in the fractures, in the surrounding rock matrix and the coupling conditions between each fracture and the corresponding portion of the rock matrix are the same as in (7), or equivalently (8), but separately for all the pieces. The reduction process, similarly to the previous part, replaces the equi-dimensional domain, which represents the intersecting region, to a single point, which is $\boldsymbol{i}_p := \bigcap_{k=1}^{n_{\mathrm{f}}} \hat{\gamma}_k$ with n_{f} the number of participating branches, and introduce a new variable $\hat{p}_I$ which represents the pressure in the intersection.

The equi-dimensional setting for a crossing is shown in Fig. 3 on the example of four intersecting fracture branches. The equi-dimensional model domain can be

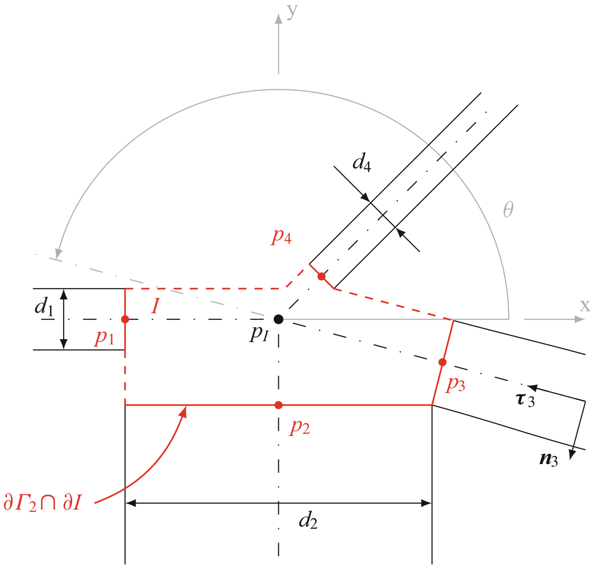


Fig. 3 Crossing with intersection geometries and location of pressure unknowns in the equidimensional model

decomposed into three different domain types: matrix, fracture and crossing, namely

$$\mathcal{D} = \Big(\bigcup_i \Omega_i\Big) \cup \Big(\bigcup_i \Gamma_i\Big) \cup I.$$

We define the crossing area I with boundaries to the fractures (solid red lines) and boundaries to the rock matrix (dashed red lines), respectively as

$$(\partial I)_{\mathrm{f},i} := \partial\Gamma_i \cap \partial I \quad \text{and} \quad (\partial I)_{\mathrm{m}} := \partial I \backslash (\bigcup_i \partial\Gamma_i).$$

Introducing the green boundaries in Fig. 4 by connecting the appropriate corners in our crossing region, we get a closed control area for which we can write the mass conservation equation which gives a relation between the crossing area pressure and the adjacent fracture pressures p_i. For the reduced model, mass conservation implies

$$\sum_{k=1}^{n_\mathrm{f}} \hat{\boldsymbol{u}}_k \cdot \boldsymbol{\tau}_k|_{\boldsymbol{i}_p} = f_I, \tag{9}$$

where $\boldsymbol{\tau}_k$ is the unit tangent along the fracture branch $\hat{\gamma}_k$, or, in other words, $\boldsymbol{\tau}_k|_{\boldsymbol{i}_p}$ is the unit outward normal of $\hat{\gamma}_k$ at $\boldsymbol{i}_p$. Moreover, $f_I \in \mathbb{R}$ is an integrated source term given for the intersection.

In the following, we will distinguish three different approaches for assigning boundary/coupling conditions for the fracture branches $\hat{\gamma}_k$ at the intersection point $\boldsymbol{i}_p$. The first two admit rather general intersection situations and are mainly suited

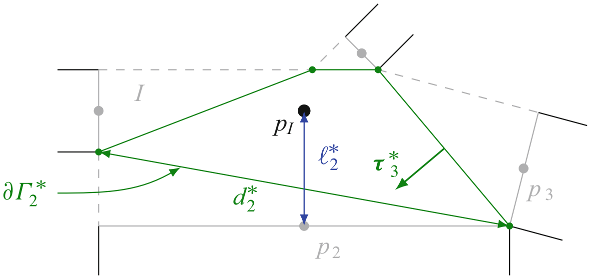


Fig. 4 Crossing with definition of geometrical parameters inside the crossing area

for the primal formulation in connection with assigning degrees of freedom in the intersection point for every fracture branch: assuming pressure continuity and Robin-type conditions. The third assumes an X-shaped intersection of four fracture branches and is especially tailored for the dual formulation in connection with a lower-dimensional XFEM approach to capture the discontinuity in the intersection.

2.3.1 Assuming Pressure Continuity

If the properties, e.g., permeability and aperture, of the fractures and in the intersection are equal, or at least comparable, a simple strategy is to impose a pressure and normal flux continuity. The model is valid also if we can consider the intersection as a void space, i.e., infinite permeability, or small enough that it can be neglected. In the latter case the source term at the intersection may be omitted. Following [4, 5, 10] and the references therein, we require mass conservation (9) together with pressure continuity

$$\hat{p}_k|_{i_p} = \hat{p}_I \quad \forall k = 1, \ldots, n_{\mathrm{f}}. \tag{10}$$

With (9) and (10), it is possible to eliminate the value $\hat{p}_I$ of the pressure at the intersection. Moreover the primal formulation of (10) is straightforward. In some cases the heterogeneity between fractures could be severe and the aforementioned model behaves poorly, see [27].

2.3.2 Robin Boundary Conditions

In [51], an alternative to requiring pressure continuity in the intersection has been proposed which amounts to replace the Dirichlet-type coupling (10) by Robin-type conditions for each fracture branch.

Considering the equi-dimensional setup from Figs. 3 and 4, we assume that the Darcy velocity $\boldsymbol{u}_k$ associated with the fracture branch Γ_k can be prolongated to the intersection region I and be defined there as

$$\boldsymbol{u}_k|_I = -\Lambda_I \boldsymbol{\tau}_k \frac{1}{\ell_k^*}(p_I - p_k),$$

where ℓ_k^* is the distance between the crossing point $\boldsymbol{i}_p$ and the point $\partial\Gamma_k \cap \hat{\gamma}_k$. Proceeding to the reduced model and integrating along the green lines in Fig. 4 yields the Robin boundary condition

$$\hat{\boldsymbol{u}}_k \cdot \boldsymbol{\tau}_k|_{\boldsymbol{i}_p} = \boldsymbol{\tau}_k^{*\top} \Lambda_I \boldsymbol{\tau}_k \frac{d_k^*}{\ell_k^*}(\hat{p}_k - \hat{p}_I), \tag{11}$$

where $d_k^* = |\partial\Gamma_k^*|$ is the length of the interface (green) for fracture k within I and $\boldsymbol{\tau}_k^*$ the unit outward normal on that interface.

The mass conservation (9) can be rewritten as

$$\sum_{k=1}^{n_\mathrm{f}} \boldsymbol{\tau}_k^{*\top} \Lambda_I \boldsymbol{\tau}_k \frac{d_k^*}{\ell_k^*}(\hat{p}_k - \hat{p}_I) = f_I. \tag{12}$$

Conditions (11) and (12) can be easily incorporated into the primal hybrid-dimensional problem formulation (8).

2.3.3 Dual Formulation for X-Shaped Intersections

In the case of two intersecting fractures with an X-shaped intersection, a mathematically rigorous derivation of coupling conditions is presented in [27, 35]. These conditions are perfectly suited for incorporation into a dual problem formulation discretized by XFEM. In this case, we have $n_\mathrm{f} = 4$ but we make an explicit use of the fact that two distinct fractures intersect and associate only one index k with the two branches of one fracture, see Fig. 5. This allows to formulate for a quantity $\hat{q}_k$ associated with fracture k its average $\{\!\{\hat{q}_k\}\!\}_{\boldsymbol{i}_p}$ and jump $[\![\hat{q}_k]\!]_{\boldsymbol{i}_p}$ at the intersection point $\boldsymbol{i}_p$. The model takes into account the aperture, permeability and angle at the intersection between fractures as well as the permeability in the intersecting region.

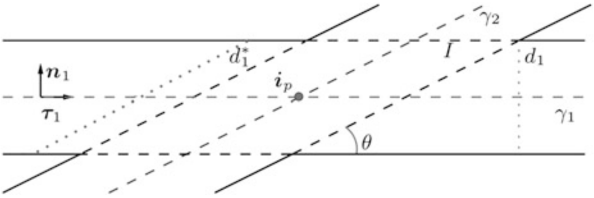


Fig. 5 Geometry and notation for the case of X-shaped intersections

This model allows a pressure and Darcy velocity discontinuity across the intersection, where the jumps are computed accounting for the pressure and fluxes from both fractures. In addition to the mass conservation (9), the coupling conditions are

$$\begin{cases} \dfrac{|I|}{d_i}\displaystyle\sum_{k=1}^{2}\dfrac{\hat{\eta}_{ik}}{d_k^*}\{\!\{\hat{\boldsymbol{u}}_k\cdot\boldsymbol{\tau}_k\}\!\}_{i_p} = [\![\hat{p}_i]\!]_{i_p} \\ \hat{\xi}_0\dfrac{d_j}{d_i}\hat{\eta}_{ii}\,[\![\hat{\boldsymbol{u}}_i\cdot\boldsymbol{\tau}_i]\!]_{i_p} = \{\!\{\hat{p}_i\}\!\}_{i_p} - \hat{p}_I \end{cases} \qquad \text{for } i,j = 1,2, i \neq j, \tag{13}$$

where $d_k^* = d_k/\sin\theta$ and θ is the angle between the two fractures, $\hat{\eta}_{ij}$ is the tangential projection along $\hat{\gamma}_j$ and then $\hat{\gamma}_i$ of the inverse of the permeability in the intersection region, namely $\hat{\eta}_{ij} := \boldsymbol{\tau}_i^\top \Lambda_I^{-1}\boldsymbol{\tau}_j$. Note the similarity of (13) to the interface conditions (7b-bis) of the "full" dual hybrid-dimensional problem. This allows to directly apply the corresponding XFEM techniques in a lower-dimensional context. Moreover, the system of equations (13) can be viewed as a generalization of (10) since the former boils down to the latter providing the intersection permeability goes to infinity or the dimension of the intersection goes to zero.

2.4 Boundary Conditions

This section is divided into two parts: Dirichlet boundary conditions for a matrix boundary that is intersected by fractures and conditions for fracture tips that are located in the interior of the matrix domain.

2.4.1 Dirichlet Conditions for Fractured Porous Media

Boundary conditions for fractured porous media systems with explicitly modeled fractures are not easy to define. The simplest choice is to prescribe a constant pressure along a domain boundary or a linear change, for example for the case of a hydrostatic pressure distribution. This often does not reflect the highly heterogeneous structure in the case of fractured porous media systems. That again leads to a strong influence of the boundary conditions on the solution if the domain is not chosen large enough. For field scale simulations, one usually obtains pointwise pressure information from which the best boundary conditions are to be picked. In [51], a possibility is presented to interpolate pointwise pressure data along a given boundary including the information of the geometrical position and geological parameters (aperture, permeability) of the fractures intersecting with this boundary.

In particular, the situation depicted in Fig. 6 is considered. For a boundary segment $\omega \subset \partial\Omega$ that is parametrized by $\chi \in [0,1]$, the left and right pressure values p_{left} and p_{right} at $\chi = 0$ and $\chi = 1$ are assumed to be known. The segment

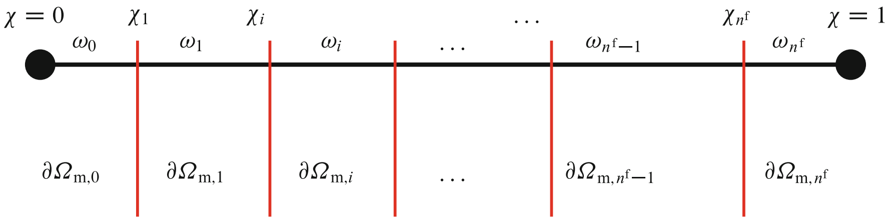


Fig. 6 Partitioning of the boundary according to intersecting fractures

ω is intersected by n^f fractures that divide it in $n^\mathrm{f} + 1$ parts ω_i. The goal is to find a pressure distribution p_b on ω that accounts for the presence of the intersecting fractures by admitting jumps across the fracture-boundary intersections χ_j. Proceeding analogously to the derivation of Problem 6, one has to find p_b such that $p_\mathrm{b}(0) = p_\mathrm{left}$, $p_\mathrm{b}(1) = p_\mathrm{right}$ and

$$\sum_{i=0}^{n^\mathrm{f}} (\lambda_{\mathrm{b},i} \nabla p_\mathrm{b}, \nabla q_\mathrm{b})_{\omega_i} + \sum_{j=1}^{n^\mathrm{f}} \lambda_{\hat{\gamma}}(\chi_j) \left([\![p_\mathrm{b}]\!]_j [\![q_\mathrm{b}]\!]_j + \xi_0 (\{\!\!\{ p_\mathrm{b} \}\!\!\}_j - p_{\mathrm{f}_j}) \{\!\!\{ q_\mathrm{b} \}\!\!\}_j \right) = 0, \tag{14}$$

for all test functions q_b. Here, $\lambda_{\mathrm{b},i} = \boldsymbol{\tau}_\mathrm{b}^\top \Lambda_i \boldsymbol{\tau}_\mathrm{b}$ is the permeability along the boundary segment, while $[\![\cdot]\!]_j$ and $\{\!\!\{ \cdot \}\!\!\}_j$ refer to the jump and average in the fracture-boundary intersection χ_j. In order to derive a closed system, one is left with the choice of the fracture pressures p_{f_j}. In [51], two options are discussed. The first one assumes $p_{\mathrm{f}_j} = \{\!\!\{ p_\mathrm{b} \}\!\!\}_j$ such that the fracture pressures are indeed an outcome of solving (14). The second one builds upon expert knowledge to describe explicit values for p_{f_j}.

In order to facilitate the solution of (14), it is assumed that p_b is piecewise linear with respect to the unfractured parts ω_i, namely,

$$p_\mathrm{b}(\chi) = m_i \chi + b_i \text{ on } \omega_i, \tag{15}$$

with the two unknown coefficients m_i, b_i. By choosing appropriate test functions q_b, analytical expressions for these coefficients are derived for an arbitrary number of fractures and the case $p_{\mathrm{f}_j} = \{\!\!\{ p_\mathrm{b} \}\!\!\}_j$, see [51].

For example, for one single fracture, $n^\mathrm{f} = 1$, the slopes are given by

$$m_0 = \frac{\mathrm{p}_\mathrm{right} - \mathrm{p}_\mathrm{left}}{\frac{\lambda_{\mathrm{b},0}}{\lambda_{\hat{\gamma}}} + |\omega_1| + |\omega_0| \frac{\lambda_{\mathrm{b},0}}{\lambda_{\mathrm{b},1}} + (\frac{\lambda_{\mathrm{b},0}}{\lambda_{\mathrm{b},1}} - 1)},$$

$$m_1 = \frac{\mathrm{p}_\mathrm{right} - \mathrm{p}_\mathrm{left}}{\frac{\lambda_{\mathrm{b},1}}{\lambda_{\hat{\gamma}}} + |\omega_0| + |\omega_1| \frac{\lambda_{\mathrm{b},1}}{\lambda_{\mathrm{b},0}} + (1 - \frac{\lambda_{\mathrm{b},1}}{\lambda_{\mathrm{b},0}})}.$$

In [51], the resulting boundary conditions are shown to be superior to standard choices. In particular, the transition from the inner part of the domain to the boundary appears much more natural in the presence of blocking fractures.

2.4.2 Boundary and Coupling Conditions for Fracture Tips

We consider now the situation where parts of the fractures are ending inside the interior of the matrix domain, namely, in $\hat{\gamma}_{\mathrm{tip}} = \partial\hat{\gamma} \cap \Omega$. Apart from being challenging from the mathematical and numerical point of view, the modeling question is what kind of boundary or coupling conditions should be prescribed at $\hat{\gamma}_{\mathrm{tip}}$.

An obvious easy choice is to prescribe no-flow conditions across the fracture tip [7], namely,

$$\hat{\boldsymbol{u}} \cdot \boldsymbol{\tau}|_{\hat{\gamma}_{\mathrm{tip}}} = 0. \tag{16}$$

In many situations, this condition is well justified by the essential modeling assumption that the fracture aperture d is small compared to its lateral dimensions. However, there can be problem settings where (16) may not be accurate enough. In particular, if the tangential permeability of the fracture is larger than the normal one, namely, $\hat{\lambda} > \lambda_{\hat{\gamma}}$, the flow across the fracture tip could be rather large compared to the flow over the fracture's lateral boundaries and should be taken into consideration.

Taking into account the flow across the tip can be achieved by assigning a corresponding source term f_Ω for the matrix domain [50],

$$f_\Omega = \delta_{\hat{\gamma}_{\mathrm{tip}}} \hat{\boldsymbol{u}} \cdot \boldsymbol{\tau}|_{\hat{\gamma}_{\mathrm{tip}}}. \tag{17}$$

The coupling condition (17) can be complemented by a condition involving the matrix and fracture pressures. For example, one could aim for

$$\hat{p}|_{\hat{\gamma}_{\mathrm{tip}}} = [\![p]\!]_{\hat{\gamma}_{\mathrm{tip}}}. \tag{18}$$

Conditions (17) and (18) are discussed and investigated for the discretized primal formulation in [50]. While (17) is implemented as a source for the matrix domain, (18) can be realized as a Dirichlet condition for the fracture. However, a proper mathematical derivation from the continuous setting as well as thorough numerical comparisons with (16) are still missing.

3 Numerical Discretization by Means of XFEM

Before we present the numerical approximation of the previous reduced models, both in dual and primal form, using the extended finite element method (XFEM), we provide a very brief overview of its historical development and some pointers to the literature.

Ideally, one would want to use a mesh that is as structured and axis-aligned as possible. The standard Galerkin finite-element method, however, cannot handle discontinuities in the solution except by resolving them through the grid, namely, by doubling and decoupling the degrees of freedom along the discontinuities. Coming from the structural-mechanics problem of evolving cracks that leads to discontinuities in the solution (displacement, stress, strain), an extension to the standard finite-element scheme was developed, [22, 23, 44], and called "eXtended Finite Element Method."

From the more theoretical point, Nitsche's method, intentionally developed to handle Dirichlet constraints, evolved to a new possibility to treat interface problems, [16, 38, 39]. XFEM and Nitsche's method applied to interface problems are in this case essentially the same approach. An overview of recent problems where XFEM methods are investigated is given in [1]. Some works that influence the following presentation are [23, 40, 44]. XFEM was first used in the fractured porous media context in [19, 33] for lower dimensional fractures introducing a discontinuous solution in the matrix, in [11] for lower dimensional fracture networks having different permeabilities in the network and therefore also discontinuities, and in [41] for thin heterogeneities (equi-dimensional) which are not resolved directly with the grid but rather with the XFEM.

In the literature such techniques are very often referred to as "partition-of-unity" PUFEM and "generalized finite-element methods" GFEM. The difference here is that those are usually on a global level where XFEM adopts the same techniques on an element-local formulation. The composite finite element method, first presented in [36], is a special type of a geometric multi-grid methods and falls therefore in the category of multi-scale methods.

3.1 Modification and Addition of Basis Functions

In the classical Galerkin finite-element approach, the discrete solution, $p_h(\boldsymbol{x})$, at a global point $\boldsymbol{x}$ in space, which lies within an element E, is defined by the sum over all shape functions associated with this element multiplied by the value of the corresponding degree of freedom $\tilde{p}_i$, cf. for example [14],

$$p_h(\boldsymbol{x}) = \sum_{i \in \mathsf{N}^E} b_i(\boldsymbol{x})\, \tilde{p}_i. \tag{19}$$

Here, b_i denotes the shape function of the degree of freedom i, $\mathsf{N}^E = \{\mathsf{n}_1, \ldots, \mathsf{n}_\mathrm{r}\}$ denotes the set of standard degrees of freedom of the element E. All matrix elements which are not cut by a fracture are treated with such a standard finite element approach.

If an element E is cut by a fracture, additional degrees of freedom $\tilde{p}_j^{\mathrm{e}}$ are introduced. Those elements which are cut by at least one fracture are called enriched elements. The discrete solution on an enriched element E can be written as

$$p_h(\boldsymbol{x}) = \sum_{i \in \mathsf{N}^E} b_i(\boldsymbol{x})\, u_i^{\mathrm{s}}(\boldsymbol{x})\, \tilde{p}_i + \sum_{j \in \mathsf{N}^{\mathrm{e},E}} b_j^{\mathrm{e}}(\boldsymbol{x})\, u_j^{\mathrm{e}}(\boldsymbol{x})\, \tilde{p}_j^{\mathrm{e}} \tag{20}$$

Here, $\mathsf{N}^{\mathrm{e},E}$ is the set of enriched degrees of freedoms. To capture discontinuities in the solution the basis functions are multiplied by discontinuous functions, where u_i^{s} denotes the discontinuity functions for the standard degrees of freedom, while u_j^{e} denotes the discontinuity functions for the enriched degrees of freedom, respectively. These functions will be defined below.

In the course of this work, the standard basis at cut elements is chosen to be same as for uncut elements, i.e., (bi-)linear and the additional shape-functions are chosen to be of the same type as the standard shape-functions, i.e., also (bi-)linear, $b_i = b_j^{\mathrm{e}}$ if i and j refer to degrees of freedom located at the same vertex. Then, b_j^{e} denotes the nodal shape function of an enriched node j. Furthermore, the XFEM concept is here used in combination with the Ritz-Galerkin approach, i.e., the basis-function space and the test-function space are equal.

The choice of the discontinuity functions is somehow arbitrary, as long as certain conditions are fulfilled. One commonly desired goal is to choose the discontinuity functions such that the resulting enriched basis functions are forced to be zero in all nodes. On the one hand this leads to a propitious quality: the nodal interpolation is still guaranteed by the solution in the standard nodes alone, [44]. More importantly this property yields to enriched basis functions which are completely local with respect to the cut elements and every basis function has only one discontinuity (within this element) for every set of additional degrees of freedom. This avoids blending elements, [28], which have to be introduced otherwise. However, this is only valid for the special case of a single interface per element. The general, more complex case of several (intersecting) interfaces is more demanding and discussed in, for example, [50]. There are many other possible choices for the discontinuity functions with different properties. For example, they can be chosen such that the standard basis remains unmodified and the discontinuity is only represented by the enriched basis or such that the mean of the enriched basis functions is zero.

To become more explicit, we define the discontinuity functions by using the sign function sgn which is positive one on the side of the positive normal direction and negative one on the other, as

$$u_i^{\mathrm{s}}(\boldsymbol{x}) := \frac{1}{2}\left|\operatorname{sgn}(\boldsymbol{x}) + \operatorname{sgn}(\boldsymbol{x}_i)\right|, \quad u_i^{\mathrm{e}}(\boldsymbol{x}) := \frac{1}{2}\left|\operatorname{sgn}(\boldsymbol{x}) - \operatorname{sgn}(\boldsymbol{x}_i)\right|.$$

The second term, $\operatorname{sgn}(\boldsymbol{x}_i)$, associates a constant value to every node, so that the discontinuity function for the original degrees of freedom is one if $\boldsymbol{x}$ and $\boldsymbol{x}_i$ lie on the same side of the interface and zero if they are on different sides, and vice versa for the additional degrees of freedom. The modified basis functions for this kind of discontinuity functions are exemplarily shown for the one-dimensional case in Fig. 7. For this approach, the orientation of the normal vector $\boldsymbol{n}$ of the interface has to be chosen. This choice is arbitrary.

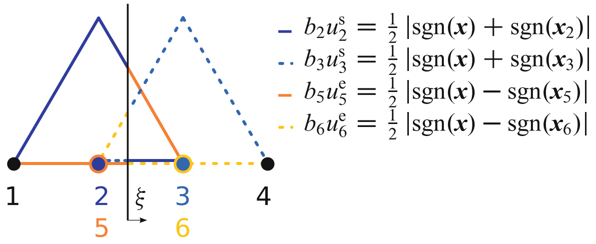


Fig. 7 The zero dimensional fracture with local coordinate ξ divides the one dimensional matrix element. *Solid lines* show the two modified basis functions associated to node two (degrees of freedom two and five), *dashed lines* the modified basis of degrees of freedom three and six at node three.

3.2 Primal Formulation with XFEM

The matrix domain Ω is discretized by n^{m} triangular or quadrilateral elements E^j into $\mathsf{T}_{\mathrm{m}}^h = \{E^j\}_{j=1}^{n^{\mathrm{m}}}$ independent of γ. The fracture γ is discretized with lower-dimensional elements, $\mathsf{T}_{\mathrm{f}}^h = \{E_{\mathrm{f}}^j\}_{j=1}^{n^{\mathrm{f}}}$, independent of $\mathsf{T}_{\mathrm{m}}^h$. We define all elements $E \in \mathsf{T}_{\mathrm{m}}^h$ which are totally in or partly belonging to Ω_i as $E_i = E \cap \Omega_i$. All elements which are not fully included in one sub-domain belong to both. The discrete space can then be defined as

$$\mathcal{Q}_i^h = \{q_{h,i} \in C^0(\Omega_i) : q_{h,i}|_{E\cap\Omega_i} \in \mathbb{Q}_1(E_i),\, E \in \mathsf{T}_{\mathrm{m}}^h\}$$

for quadrilateral elements E, with $\mathbb{Q}_1$ being replaced by $\mathbb{P}_1$ for triangular elements. The complete discrete space for the rock matrix domain is then just the product space of the sub-domain spaces $\mathcal{Q}_{\mathrm{m}}^h = \mathcal{Q}_1^h \times \mathcal{Q}_2^h$, where the elements cut by a fracture are contained twice but each with the cut basis. The discrete space for γ reads

$$\mathcal{Q}_{\mathrm{f}}^h = \{\hat{q}_h \in C^0(\gamma) : \hat{q}_h|_{E_{\mathrm{f}}} \in \mathbb{Q}_1(E_{\mathrm{f}}), E_{\mathrm{f}} \in \mathsf{T}_{\mathrm{f}}^h\}$$

so that the combined space is $\mathcal{Q}^h = \mathcal{Q}_{\mathrm{m}}^h \times \mathcal{Q}_{\mathrm{f}}^h$. This allows to obtain the discrete formulation of Problem 6.

Problem 7 (Primal Hybrid-Dimensional Discrete Formulation) Find $p_h = (p_{h,1}, p_{h,2})$ and $\hat{p}_h$ in subspaces of $\mathcal{Q}_{\mathrm{m}}^h$ and $\mathcal{Q}_{\mathrm{f}}^h$ that respect the given boundary conditions such that

$$\begin{cases} a_{\mathrm{p}}(p_h, q_h) - c_{\mathrm{p}}(q_h, \hat{p}_h) = F(q_h) \\ -c_{\mathrm{p}}(p_h, \hat{q}_h) + \hat{a}_{\mathrm{p}}(\hat{p}_h, \hat{q}_h) = \hat{F}(\hat{q}_h) \end{cases},$$

for all test functions q_h and $\hat{q}_h$ defined in proper subspaces of $\mathcal{Q}_{\mathrm{m}}^h$ and $\mathcal{Q}_{\mathrm{f}}^h$.

3.3 Dual Mixed Formulation

In [19], and in some more recent works such as [30–32] the XFEM is applied to the dual mixed formulation of the problem in a similar way, but with different FEM spaces. In particular, the lowest order Raviart-Thomas pair $\mathbb{RT}_0, \mathbb{P}_0$, see [47, 49], is employed for velocity and pressure, respectively. This is a common choice in porous media simulations, which guarantees local mass conservation. In the aforementioned works the domain is discretized by means of a triangular or tetrahedral grid, however, the method could be generalized to the case of quadrilateral or hexahedral grids.

The cut mixed finite element spaces can be defined as follows. For each element E_m let $\mathbb{RT}_0(E_{\mathrm{m},i}) = \left\{\mathbf{v}_h|_{E_{\mathrm{m},i}} : \mathbf{v}_h \in \mathbb{RT}_0(E_\mathrm{m})\right\}$ be the restriction of the standard $\mathbb{RT}_0$ functions to the sub-element $E_{\mathrm{m},i}$, and analogously let $\mathbb{P}_0(E_{\mathrm{m},i}) = \left\{q_h|_{E_{\mathrm{m},i}} : q_h \in \mathbb{P}_0(E_\mathrm{m})\right\}$ be the restriction of the standard $\mathbb{P}_0$ functions. See Fig. 8 for a sketch of the restricted basis functions and the corresponding degrees of freedom in the 2D case.

The discrete velocities and pressure in Ω are then sought in the following spaces respectively:

$$\mathbf{V}_h = \mathbf{V}_{1,h} \times \mathbf{V}_{2,h} \quad Q_h = Q_{1,h} \times Q_{2,h}$$

where

$$\mathbf{V}_{i,h} = \left\{\mathbf{v}_h \in \mathbf{H}_{div}(\Omega_i) : \mathbf{v}_h \in \mathbb{RT}_0(E_{\mathrm{m},i}) \ \forall E_\mathrm{m} \in \mathsf{T}_\mathrm{m}^h\right\}$$

$$Q_{i,h} = \left\{q_h \in L^2(\Omega_i) : q_h \in \mathbb{P}_0(E_{\mathrm{m},i}) \ \forall E_\mathrm{m} \in \mathsf{T}_\mathrm{m}^h\right\}.$$

The finite element spaces for the fracture problem, on γ, are the standard $\mathbb{RT}_0$-$\mathbb{P}_0$ in $N-1$ dimensions, thus, the discrete flux and pressure in the fracture are sought in the spaces

$$\hat{\mathbf{V}}_{i,h} = \left\{\hat{\mathbf{v}}_h \in \mathbf{H}_{div}(\gamma) : \hat{\mathbf{v}}_h \in \mathbb{RT}_0(E_\mathrm{f}) \ \forall E_\mathrm{f} \in \mathsf{T}_\mathrm{f}^h\right\}$$

$$\hat{Q}_{i,h} = \left\{\hat{q}_h \in L^2(\gamma) : \hat{q}_h \in \mathbb{P}_0(E_\mathrm{f}) \ \forall E_\mathrm{f} \in \mathsf{T}_\mathrm{f}^h\right\}.$$

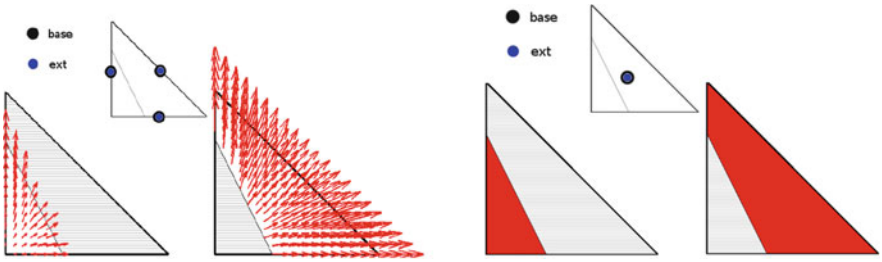


Fig. 8 Basis functions for the lowest order Raviart-Thomas pair, restricted to the subdomains Ω_i

In the case of intersecting fractures, one could consider non-matching fracture grids at the intersection: in this case, a suitable XFEM enrichment should be considered also in the fractures. We refer to [27] for details.

We can now define $\tilde{\mathbf{V}}_h = \mathbf{V}_h \times \hat{\mathbf{V}}$, and $\tilde{Q}_h = Q_h \times \widehat{Q_h}$, and $\mathbf{W}_h = \tilde{\mathbf{V}}_h \times \tilde{Q}_h$ and formulate the discrete version of Problem 4.

Problem 8 (Dual Hybrid-Dimensional Discrete Formulation) Find $(\mathbf{u}_h, \hat{u}_h, p_h, \hat{p}_h) \in \mathbf{W}_h$ such that

$$a_h(\mathbf{u}_h, \hat{\mathbf{u}}_h, \mathbf{v}_h, \hat{\mathbf{v}}_h) + b_h(p_h, \hat{p}_h, \mathbf{v}_h, \hat{\mathbf{v}}_h) - b_h(q_h, \hat{q}_h, \mathbf{u}_h, \hat{\mathbf{u}}_h) =$$

$$\mathcal{F}(\mathbf{v}_h, \hat{\mathbf{v}}_h, q_h, \hat{q}_h) \; \forall (\mathbf{v}_h, \hat{\mathbf{v}}_h, q_h, \hat{q}_h) \in \mathbf{W}_h$$

for all test functions q_h and $\hat{q}_h$ defined in proper subspaces of $\mathcal{Q}^h_{\mathrm{m}}$ and $\mathcal{Q}^h_{\mathrm{f}}$.

The well-posedness of the dual discrete problem has been proven in [19] for the case of given pressure in the fracture, and in [21] for the fully coupled case. Particularly relevant is the problem of the inf-sup stability of the extended spaces: indeed, even if we start from a stable pair the enriched spaces could present instabilities in some particular configurations. Sufficient conditions on the fracture geometry and on the underlying grid of the porous medium are given in the two aforementioned works.

3.4 Fracture Grids and Approximation of the Coupling Terms

Once the finite element spaces for both the fracture and the rock matrix are defined, one crucial ingredient is the approximation of the coupling term between the two media. To simplify the presentation, we consider only fracture and matrix grids which are *genuinely* non-matching if each fracture element is fully contained in a N-dimensional element or is contained in a pair of facing N-dimensional elements. In the important case of matching geometries, a standard technique can be employed to approximate the coupling term, see for example [29]. The construction of the fracture grid can be done in two different ways. One possibility is to consider the fracture grid induced by the intersection between the background mesh and the interface. This approach avoids the construction of complex interpolation operators, explained in the sequel, but can be done easily only in the two-dimensional case and may produce elements with strongly varying aspect ratios. One possibility to overcome these difficulties consists in using the trace of higher dimensional basis functions as in the Trace FEM method, coupled with suitable stabilizations, see [17, 43, 45]. To allow a higher flexibility in the numerical discretization, it is possible to introduce an interpolation operator $\mathcal{M} : \hat{Q}_h \to Q_h$ which maps the value of the pressure in the fracture elements to the corresponding element in the matrix grid. Note that it has to take into account also fractions of fracture elements. In the particular case of a piecewise constant approximation for pressure, following [34], the discrete version of $\mathcal{M}$ is a rectangular matrix $\mathcal{M}_h$ with entries

$$[\mathcal{M}_h]_{ij} = \int_{\hat{K}_j \cap K_i} 1 \mathrm{d}x,$$

where K_i is the i-th element in matrix mesh and $\hat{K}$ is the j-th element in the fracture mesh. To preserve mass conservation at discrete level, we consider the approximation of the adjoint operator $\mathcal{M}^* : Q_h \to \hat{Q}_h$ as the transpose of $\mathcal{M}_h$.

In [29], the authors note that, in the case of an immersed fracture, if the fracture grid is too fine compared with the mesh of the matrix, depending also on the permeability contrast, the solution in the fracture could present oscillations. The authors suggest a possible explanation which is related to the singularity of the solution at the fracture tip. The aforementioned work is in the context of non-matching, but still conforming (i.e. aligned with the grid) discretizations, and the authors consider a mortar technique to deal with the non-matching elements at the interface. The same phenomenon is also observed when the XFEM method is employed. In fact considering the simple domain depicted in Fig. 9, it is possible to obtain the solutions reported in Fig. 10 where the fracture behaves like a barrier only in its middle part. The domain is discretized using a structured triangular mesh where each boundary edge is approximated with n segments, while the fracture is discretized using m segments. The results show that, also in the case of XFEM, the oscillations occur when the discretization of the fracture is finer than the rock mesh. Apart from cases where the mesh is too coarse the oscillations exhibit a frequency that depends only on the discretization of the outer medium. In this particular case it can be estimated as $n/2$ for most of the cases. We notice that this frequency corresponds to the $n-$th eigenfunction of the problem. Indeed, if n is odd the solution becomes asymmetrical. It is also interesting that, for fixed n, the amplitude of the oscillations is constant with increasing m, while if we refine both fracture and

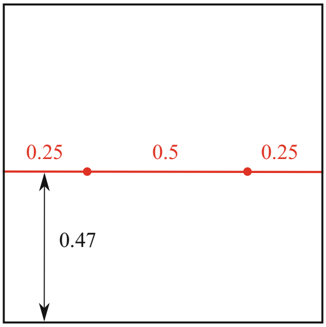


Fig. 9 Representation of the domain for the oscillation problem. In the fracture we highlight the pieces with different permeability

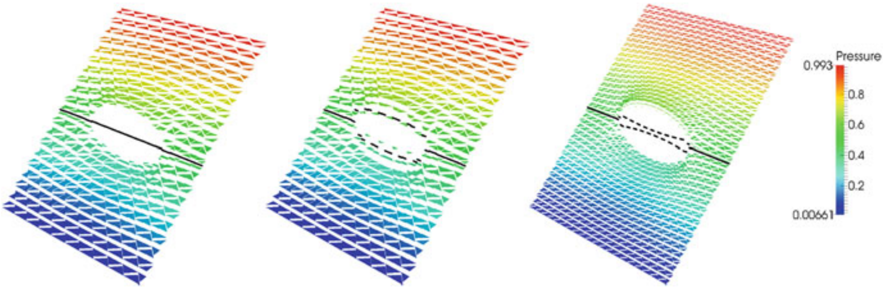


Fig. 10 In the *left*, the grid sizes of the matrix and the fracture are comparable. In the *centre*, the fracture grid is finer than the rock grid, some oscillations are present. In the *right*, both meshes are refined maintaining the same ratio of the grid size as the solution in the *centre*. In this case the amplitude of the oscillations decreases

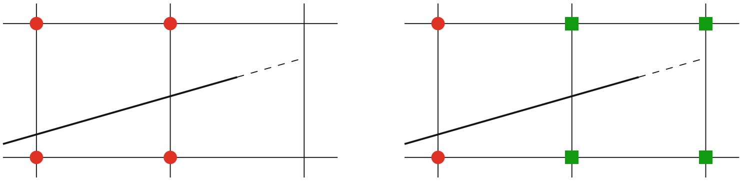


Fig. 11 Basis function enrichment around fracture tips. *Red circles* indicate XFEM enrichment, *green squares* enrichment by radial functions. *Left*: pure XFEM enrichment according to [50]. *Right*: XFEM and radial functions enrichment as suggested in [37]

medium the amplitude decreases. This phenomena are particularly relevant when the normal or tangential (or both) permeability change sharply along the fracture.

3.5 *Basis Function Enrichment Around Fracture Tips*

If a fracture branch ends inside the interior of the matrix domain in the fracture tip $\hat{\gamma}_{\text{tip}}$, the question arises on how to enrich the basis functions inside an element E_{tip} that contains $\hat{\gamma}_{\text{tip}}$. For the primal formulation, an ad-hoc solution is presented and used in [50]. In E_{tip}, only those basis functions are enriched that correspond to the vertices of the element face that is intersected by the fracture, see the left picture of Fig. 11. While this approach is attractive for its simplicity, it disregards the potential singularity in the solution at the tip $\hat{\gamma}_{\text{tip}}$. As described in [37], it is more appropriate to account explicitly for this singularity by adding radial functions

$$g_1(r,\theta) = \sqrt{r}\sin(\theta/2), \quad g_2(r,\theta) = \sqrt{r}\cos(\theta/2),$$

multiplied with the standard basis functions b_i in E_{tip}. Here, r is the distance from $\hat{\gamma}_{\text{tip}}$, while θ indicates the angle with respect to the tangential fracture direction. This situation is depicted in the right picture of Fig. 11.

4 Solvers

In this section, we discuss some issues related with the numerical solution of the system resulting from numerical discretization. In particular, we present the problem of ill-conditioning in the presence of very small sub-elements, and the possibility of using iterative strategies in a domain-decomposition framework as an alternative to a fully monolithic approach.

4.1 *Conditioning*

Since we are considering an arbitrary position of the interface with respect to the underlying grid, it is possible that, when an element E is cut into two sub-elements E_1, E_2, the ratio $\frac{|E_i|}{|E|}$ is very small or even zero for some elements/some configurations of the interface. The case of $|E_i| = 0$, corresponding to a fracture that contains one or two adjacent vertices of the element, is an exception ruled out by the assumptions of the XFEM method and should be handled separately, while the case of small sub-elements can lead to ill-conditioned matrices. In particular, the condition number can degenerate as $\mu(\mathsf{T}^h_{\mathrm{m}}, \hat{\gamma}) = \min_{E\in\mathcal{G}} \min_{i=1,2} \frac{|E_i|}{|E|}$ tends to zero, where $\mathcal{G}$ denotes the set of elements that are cut by the interface. This problem has been investigated for the case of a mixed formulation in [19]. In this work, a numerical experiment is presented, where, changing the position of a vertical fracture on a structured grid, the authors obtain smaller and smaller values of $\mu(\mathsf{T}^h_{\mathrm{m}}, \hat{\gamma})$ and compute the corresponding maximum and minimum eigenvalue of the matrix C defined as

$$C = \begin{bmatrix} A & B^T \\ -B & 0 \end{bmatrix},$$

arising from the discretization with XFEM of the Darcy problem in the cut bulk medium. While the maximum eigenvalue is approximately constant, the minimum decreases with μ. Even if the matrix is not symmetric and positive definite, the ratio between maximum and minimum eigenvalue

$$ic(C) = \frac{\max_i |\lambda_i(C)|}{\min_i |\lambda_i(C)|}$$

can be used as an indicator of ill-conditioning. In [19] an optimal preconditioner P for the problem is presented and tested on the same problem, showing that $ic(P^{-1}C)$ is constant for all μ. Moreover, a simpler diagonal preconditioner P_L can be used to perform matrix equilibration: in this case the preconditioner is not optimal with respect to the grid size h, but for a fixed h one still obtains a constant ratio $ic(P_L^{-1}C)$.

4.2 Iterative Approaches

The coupled problem of flow in a porous matrix and a fracture has been interpreted in a domain-decomposition framework in [2] under the assumptions of pressure continuity across the fracture, i.e. the assumptions of permeable fractures. It is shown how, in the cases of a fracture that cuts the domain Ω in two disjoint parts Ω_i the problem can be recast as a global equation on the interface $\hat{\gamma}$ for the unknown $\hat{p}$. In a more recent work, [42], the concept is generalized to the case of fractures with arbitrary permeability, including the case of (nearly) impermeable interfaces. Once again the problem can be formulated as a positive definite problem on the interface $\hat{\gamma}$. In particular, in the case $\xi = 1$ the problem simplifies again to a problem for only one scalar unknown $\hat{p}$,

$$\overline{S}_1(\hat{p}, f_1) + \overline{S}_2(\hat{p}, f_2) + \nabla_\tau \cdot (\hat{\lambda} \nabla_\tau \hat{p}) = f_f \tag{21}$$

where $\overline{S}_{1,2}(\hat{p}, f_{1,2})$ are the Robin-to-Neumann operators accounting for the coupling with the flow problem in $\Omega_{1,2}$. The interface Eq. (21) can then be solved iteratively. This approach is meant to provide an efficient method for the solution of the coupled fracture-medium problem: indeed, when considering a mixed formulation, the system is not positive definite and in realistic configurations it can be very large. Therefore, it can be convenient to eliminate some of the unknowns to obtain a problem that is easier to solve.

In the context of non-matching discretizations, an iterative approach for the solution of the coupled problem is also discussed in [19], where a similar approach is used, adding more information and numerical evidence about the convergence of the procedure. In particular, it is proven that the iterative method converges for some values of a relaxation parameter ω that must satisfy

$$\omega \gtrsim M_{\hat{\gamma}}^2 \frac{\lambda_{\hat{\gamma}}}{\xi_0 \hat{\lambda}}, \quad \text{where} \quad M_{\hat{\gamma}} = \min\left\{1, h\sqrt{\frac{\max\{\lambda_1, \lambda_2\}}{\lambda_{\hat{\gamma}}}}\right\}.$$

Numerical experiments show that, for small $\lambda_{\hat{\gamma}}/\hat{\lambda}$ ratios the relaxation parameter can be very small and convergence is achieved in just few iterations, while for higher ratios the iterative method converges very slowly.

In the relevant case of a discrete fracture network approximation, namely, where the rock matrix is supposed to be impervious, an interesting approach is proposed in [11, 12, and references therein] to solve in an efficient way the global system of equations. The authors consider the continuous coupling conditions among the fractures, as presented in Sect. 2.3.1, and a primal formulation of the problem. In this part, to simplify the notation we assume that the equi-dimensional domain are the fractures while the one-codimensional domain are the intersections among the fractures. To decouple the solution on each fracture, the transmission conditions (10) are imposed in a weak way through an optimization problem, i.e. naming S a strip of intersecting elements between two fractures solve

$$\begin{cases} \min J(p) = \min \sum\limits_{S \in \mathcal{S}} \left\| [\![p]\!]_S \right\|^2 + 2\|\{\!\{\boldsymbol{u} \cdot \boldsymbol{n}\}\!\}_S\|^2 \\ s.t. \;\; p \text{ solution in each fracture} \end{cases}$$

where $\mathcal{S}$ is the set of all the intersection regions, the norms are defined on proper spaces and $\boldsymbol{u} \cdot \boldsymbol{n}$ is a suitable reconstruction of the normal flux at the intersection S. The optimal solution of the minimum problem gives $J(p) = 0$ and can be computed numerically using a gradient method. With this method the linear system, which couple all the fractures, has much smaller size than the aforementioned approaches and the computation of the pressure in each fracture is completely parallelizable.

5 Conclusions

In this review paper, we presented several mathematical models and numerical algorithms to simulate single-phase flow in a porous medium containing fractures. Two main challenges are addressed. First, the fractures play a crucial role in subsurface flows and should be carefully accounted for to achieve reliable simulations, however their geometrical and geological data pose several difficulties from a modeling and discrete point of view. A common approach which is broadly used in the literature is to consider an hybrid-dimensional model where the fractures are treated as objects of lower dimension. Second, the position of the fractures may be unknown and several scenarios are needed to obtain a representative solution of the problem, or to speed up the simulations: for this reason it is better to allow these fractures to be geometrically decoupled from the surrounding porous medium. In this case, an XFEM approach can be a valuable option to overcome this requests. Several geometrical difficulties have been addressed in this paper, such as the treatment of intersections and tips. To the best of our knowledge, several interesting issues are still open for further investigation, such as a full three-dimensional setting for general networks of fractures with XFEM, a physical derivation of a more appropriate condition at the fracture tip, suitable stabilizations to increase the robustness of the linear solvers and a deep analysis for the case of vanishing aperture.

References

1. Abdelaziz, Y., Hamouine, A.: A survey of the extended finite element. Comput. Struct. **86**(11–12), 1141–1151 (2008)
2. Alboin, C., Jaffré, J., Roberts, J.E., Wang, X., Serres, C.: Domain decomposition for some transmission problems in flow in porous media. In: Numerical Treatment of Multiphase Flows in Porous Media (Beijing, 1999). Lecture Notes in Physics, vol. 552, pp. 22–34. Springer, Berlin (2000)

3. Alboin, C., Jaffré, J., Roberts, J.E., Serres, C.: Modeling fractures as interfaces for flow and transport in porous media. In: Fluid Flow and Transport in Porous Media, Mathematical and Numerical Treatment: Proceedings of an AMS-IMS-SIAM Joint Summer Research Conference on Fluid Flow and Transport in Porous Media, Mathematical and Numerical Treatment, Mount Holyoke College, South Hadley, Massachusetts, 17–21 June 2001, vol. 295, pp. 13–25. American Mathematical Society, Providence (2002)
4. Alboin, C., Jaffré, J., Roberts, J.E., Serres, C.: Modeling fractures as interfaces for flow and transport in porous media. In: Fluid Flow and Transport in Porous Media: Mathematical and Numerical Treatment (South Hadley, MA, 2001). Contemporary Mathematics, vol. 295, pp. 13–24. American Mathematical Society, Providence (2002)
5. Amir, L., Kern, M., Martin, V., Roberts, J.E.: Décomposition de domaine et préconditionnement pour un modèle 3D en milieu poreux fracturé. In: Proceeding of JANO 8, 8th Conference on Numerical Analysis and Optimization (2005)
6. Angot, P.: A model of fracture for elliptic problems with flux and solution jumps. C. R. Math. **337**(6), 425–430 (2003)
7. Angot, P., Boyer, F., Hubert, F.: Asymptotic and numerical modelling of flows in fractured porous media. M2AN Math. Model. Numer. Anal. **43**(2), 239–275 (2009)
8. Antonietti, P.F., Formaggia, L., Scotti, A., Verani, M., Verzotti, N.: Mimetic finite difference approximation of flows in fractured porous media. Technical Report, Politecnico di Milano (2015)
9. Bear, J.: Dynamics of Fluids in Porous Media. American Elsevier, New York (1972)
10. Benedetto, M.F., Berrone, S., Pieraccini, S., Scialò, S.: The virtual element method for discrete fracture network simulations. Comput. Methods Appl. Mech. Eng. **280**(0), 135–156 (2014)
11. Berrone, S., Pieraccini, S., Scialò, S.: On simulations of discrete fracture network flows with an optimization-based extended finite element method. SIAM J. Sci. Comput. **35**(2), 908–935 (2013)
12. Berrone, S., Pieraccini, S., Scialò, S.: An optimization approach for large scale simulations of discrete fracture network flows. J. Comput. Phys. **256**(0), 838–853 (2014)
13. Bonn, W.M., Nordbotten, J.M.: Robust discretization of flow in fractured porous media. arXiv:1601.06977 [math.NA] (2016)
14. Braess, D.: Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics. Cambridge University Press, Cambridge (2007)
15. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Computational Mathematics, vol. 15. Springer, Berlin (1991)
16. Burman, E., Hansbo, P.: Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method. Appl. Numer. Math. **62**(4), 328–341 (2012)
17. Burman, E., Claus, S., Hansbo, P., Larson, M.G., Massing, A.: CutFEM: discretizing geometry and partial differential equations. Int. J. Numer. Methods Eng. **104**(7), 472–501 (2015)
18. D'Angelo, C., Scotti, A.: A Mixed Finite Element Method for Darcy Flow in Fractured Porous Media with Non-Matching Grids. Technical Report, MOX, Mathematical Department, Politecnico di Milano, 2010
19. D'Angelo, C., Scotti, A.: A mixed finite element method for Darcy flow in fractured porous media with non-matching grids. Math. Model. Numer. Anal. **46**(02), 465–489 (2012)
20. Dassi, F., Perotto, S., Formaggia, L., Ruffo, P.: Efficient geometric reconstruction of complex geological structures. Math. Comput. Simul. **46**(02), 465–489 (2014)
21. Del Pra, M., Fumagalli, A., Scotti, A.: Well posedness of fully coupled fracture/bulk Darcy flow with XFEM. Technical Report 25/2015, Politecnico di Milano (2015). Submitted to: SIAM Journal on Numerical Analysis
22. Dolbow, J.: An extended finite element method with discontinuous enrichment for applied mechanics. Ph.D. thesis, Northwestern University (1999)
23. Dolbow, J., Moës, N., Belytschko, T.: Discontinuous enrichment in finite elements with a partition of unity method. Finite Elem. Anal. Des. **36**(3–4), 235–260 (2000)
24. Ern, A., Guermond, J.-L.: Theory and Practice of Finite Elements. Applied Mathematical Sciences. Springer, New York (2004)

25. Faille, I., Flauraud, E., Nataf, F., Pégaz-Fiornet, S., Schneider, F., Willien, F.: A new fault model in geological basin modelling. Application of finite volume scheme and domain decomposition methods. In: Finite Volumes for Complex Applications, III (Porquerolles, 2002), pp. 529–536. Hermes Science Publishing, Paris (2002)
26. Faille, I., Fumagalli, A., Jaffré, J., Roberts, J.E.: A double-layer reduced model for fault flow on slipping domains with hybrid finite volume scheme. SIAM: J. Sci. Comput. (2015, in preparation). hal-01162048
27. Formaggia, L., Fumagalli, A., Scotti, A., Ruffo, P.: A reduced model for Darcy's problem in networks of fractures. ESAIM: Math. Model. Numer. Anal. **48**, 1089–1116 (2014)
28. Fries, T.: A corrected XFEM approximation without problems in blending elements. Int. J. Numer. Methods Eng. **75**(5), 503–532 (2008)
29. Frih, N., Martin, V., Roberts, J.E., Saâda, A.: Modeling fractures as interfaces with nonmatching grids. Comput. Geosci. **16**(4), 1043–1060 (2012)
30. Fumagalli, A.: Numerical Modelling of Flows in Fractured Porous Media by the XFEM Method. Ph.D. thesis, Politecnico di Milano (2012)
31. Fumagalli, A., Scotti, A.: Numerical modelling of multiphase subsurface flow in the presence of fractures. Commun. Appl. Ind. Math. **3**(1) (2011)
32. Fumagalli, A., Scotti, A.: A reduced model for flow and transport in fractured porous media with non-matching grids. In: Proceedings of ENUMATH 2011, the 9th European Conference on Numerical Mathematics and Advanced Applications. Springer, Berlin (2012)
33. Fumagalli, A., Scotti, A.: An efficient XFEM approximation of Darcy flows in fractured porous media. MOX Report 53 (2012)
34. Fumagalli, A., Scotti, A.: A numerical method for two-phase flow in fractured porous media with non-matching grids. Adv. Water Resour. **62**(Part C(0)), 454–464 (2013). Computational Methods in Geologic CO2 Sequestration
35. Fumagalli, A., Scotti, A.: An efficient XFEM approximation of Darcy flow in arbitrarly fractured porous media. Oil Gas Sci. Technol. - Rev. d'IFP Energies Nouv. **69**(4), 555–564 (2014)
36. Hackbusch, W., Sauter, S.A.: Composite finite elements for the approximation of PDEs on domains with complicated micro-structures. Numer. Math. **75**(4), 447–472 (1997)
37. Hanowski, K., Sander, O.: Simulation of deformation and flow in fractured, poroelastic materials (2016). Preprint, arXiv:1606.05765
38. Hansbo, P.: Nitsche's method for interface problems in computational mechanics. GAMM-Mitteilungen **28**(2), 183–206 (2005)
39. Hansbo, A., Hansbo, P.: An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. Comput. Methods Appl. Mech. Eng. **191**(47–48), 5537–5552 (2002)
40. Hansbo, A., Hansbo, P.: A finite element method for the simulation of strong and weak discontinuities in solid mechanics. Comput. Methods Appl. Mech. Eng. **193**(33), 3523–3540 (2004)
41. Huang, H., Long, T.A., Wan, J., Brown, W.P.: On the use of enriched finite element method to model subsurface features in porous media flow problems. Comput. Geosci. **15**(4), 721–736 (2011)
42. Martin, V., Jaffré, J., Roberts, J.E.: Modeling fractures and barriers as interfaces for flow in porous media. SIAM J. Sci. Comput. **26**(5), 1667–1691 (2005)
43. Massing, A., Larson, M.G., Logg, A.: Efficient implementation of finite element methods on non-matching and overlapping meshes in 3d. arXiv preprint arXiv:1210.7076 (2012)
44. Mohammadi, S.: Extended Finite Element Method. Wiley, New York (2008)
45. Olshanskii M.A., Reusken, A., Grande, J.: A finite element method for elliptic equations on surfaces. SIAM J. Numer. Anal. **47**(5), 3339–3358 (2009)
46. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations. Springer Series in Computational Mathematics, vol. 23. Springer, Berlin (1994)
47. Raviart, P.-A., Thomas, J.-M.: A mixed finite element method for second order elliptic problems. Lect. Notes Math. **606**, 292–315 (1977)

48. Remij, E., Remmers, J., Huyghe, J., Smeulders, D.: The enhanced local pressure model for the accurate analysis of fluid pressure driven fracture in porous materials. Comput. Methods Appl. Mech. Eng. **286**, 296–312 (2015)
49. Roberts, J.E., Thomas, J.-M.: Mixed and Hybrid Methods. Handbook of Numerical Analysis, vol. II, pp. 523–639. North-Holland, Amsterdam (1991)
50. Schwenck, N.: An XFEM-based model for fluid flow in fractured porous media. Ph.D. thesis, Department of Hydromechanics and Modelling of Hydrosystems, University of Stuttgart (2015)
51. Schwenck, N., Flemisch, B., Helmig, R., Wohlmuth, B.: Dimensionally reduced flow models in fractured porous media: crossings and boundaries. Comput. Geosci. **19**(6), 1219–1230 (2015)
52. Tunc, X., Faille, I., Gallouët, T., Cacas, M.C., Havé, P.: A model for conductive faults with non-matching grids. Comput. Geosci. **16**, 277–296 (2012)
53. Zimmerman, R.W., Kumar, S., Bodvarsson, G.S.: Lubrication theory analysis of the permeability of rough-walled fractures. Int. J. Rock Mech. Mining Sci. Geomech. Abstr. **28**(4), 325–331 (1991)

Part II

Enhanced Finite Element Formulations for Fracture and Interface Problems

Modeling of Fracture in Polycrystalline Materials

Steffen Beese, Stefan Loehnert, and Peter Wriggers

Abstract Predicting the behaviour of fracture processes within polycrystalline microstructures will help to develop more accurate mesoscale material models and will give insight to effects which can only be measured ex-situ. Therefore a non-local damage model is introduced and coupled to finite deformation crystal plasticity. Cracks are represented sharply by using the extended finite element method in combination with level set techniques. As damage evolves cracks start to propagate. A new crack propagation algorithm is presented and studied by academic examples.

1 Introduction

In industrial forming processes metals are subjected to large plastic deformations. Related to the plastic deformation within the material microstructure pores start to grow, merge and finally lead to microcracks within the material. Macroscopically this is recognised as a loss of stiffness of the structural behaviour and a reduced strength. This effect is usually modelled with continuum damage mechanics as done by Tvergaard and Needleman in [29]. In the context of finite elements, during the last decades several theories were established to circumvent the mesh dependent localization. Most of them, see e.g. [10, 23] or [26], are based on the introduction of a new degree of freedom for the thermodynamic driving force of the damage. This contributions follows the work of Reusch et al. [26] who determinate the new degree of freedom by solving an additional scalar balance equation. Nevertheless, if damage evolves and is used to predict the initiation or propagation of cracks the global stiffness matrix becomes ill-conditioned. Furthermore cracks are just represented in a smeared way and their dimension is related to the mesh size and some artificial internal length parameter.

S. Beese (✉) • S. Loehnert • P. Wriggers
Institute of Continuum Mechanics, Leibniz Universität Hannover, Appelstraße 11, Hannover, Germany
e-mail: beese@ikm.uni-hannover.de; loehnert@ikm.uni-hannover.de; wriggers@ikm.uni-hannover.de

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_4

To overcome this drawback the damage is transferred to discrete cracks if it exceeds a material depended threshold value. The representation of fracture is modeled with the eXtended Finite Element Method (XFEM) according to [4]. In combination with the level set techniques [5] this numerical tool enables a nearly mesh independent crack representation. The position of the crack is described implicitly by level set functions, and the discontinuity within the displacement field is captured using enrichment functions and additional degrees of freedom. Since classical crack propagation criteria lose their validity in a finite deformations context and for inelastic material behaviour an alternative approach is described. The argumentation for the chosen damage based criterion is that once the crack is initiated it will propagate if the pores around the crack front will continue to coalesce. So the mechanisms for crack initiation and propagation are the same. A very elegant way to construct the crack extension implicitly is the global crack tracking algorithm proposed in Oliver and Huespe [22] and the level set update procedures proposed by Gravouil et al. [12] and Moës et al. [19]. Here a combination of both approaches is presented and a different solution method for the level set update is suggested. The accuracy of both methods is studied in a similar way Duflot proposed it in [8].

After the reconstruction of the level set they have to be re-initialised and re-orthogonalised by the solution of a Hamilton-Jacobi equation. After the propagation no remeshing is needed, but the enrichment scheme has to be updated in the vicinity of the crack front. There is also a need for adapting the history variables in these elements. This becomes necessary because the integration points are determined by a subcell integration strategy as commonly used in the XFEM since the early work [18]. Within this strategy the finite element is divided along the interface and the integration points are placed in these subcells in order to consider the discontinuous character of the displacement field. This framework is used in combination with a crystal plasticity model with viscous regularization as proposed e.g. by Steinmann and Stein [27] but within the context of elastic isomorphism [6, 28] to predict the evolution of microcracks.

2 Problem Statement

In this section the boundary value problem of a fractured body is elaborated. After a short review of the balance equation the constitutive framework is presented. By coupling of the constitutive equations for crystal plasticity with a non-local damage formulation a natural way of degrading material strength is introduced.

2.1 Balance of Linear Momentum for a Fractured Body

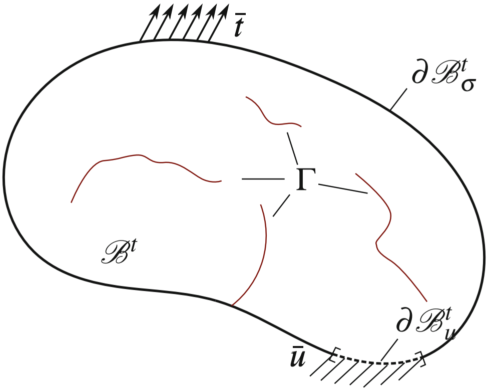
A body $\mathscr{B} \subset \mathbb{E}^3$ is a subset of the three dimensional Euclidean space and can be identified by the position $\boldsymbol{X} \in \mathbb{E}^3$ of its material points. At time $t = 0$ the body is in its initial configuration $\chi(\mathscr{B}, 0) = \mathscr{B}^0$ and at current time t the deformed body occupies the current configuration $\chi(\mathscr{B}, t) = \mathscr{B}^t$. The deformation χ in this context is a bijective mapping $\chi : \mathbb{E}^3 \times \mathbb{R} \to \mathbb{E}^3$. In Fig. 1 a fractured body $\mathscr{B}$ is schematically displayed in the current configuration. Cracks are allowed to be fully embedded in the domain and to penetrate the surface $\partial\mathscr{B}$ of the body. The crack surfaces Γ are assumed to be traction free. A surface traction $\bar{\boldsymbol{t}}$ is applied on the boundary $\partial\mathscr{B}^t_\sigma$ as Neumann condition and some displacement $\bar{\boldsymbol{u}}$ is specified on $\partial\mathscr{B}^t_u$ as Dirichlet boundary condition. If body forces are not present the strong form of the resulting quasi static boundary value problem defined in the initial configuration reads:

BVP 1 *Find* $\boldsymbol{u}(\boldsymbol{X}, t) \in \mathbb{E}^3$ *such that:*

$$\begin{aligned} \mathrm{Div}\boldsymbol{P} &= 0 && \forall \boldsymbol{X} \in \mathscr{B}^0 \\ \boldsymbol{P}\boldsymbol{N} &= \bar{\boldsymbol{T}} && \forall \boldsymbol{X} \in \partial\mathscr{B}^0_\sigma \\ \boldsymbol{P}\boldsymbol{N} &= \boldsymbol{0} && \forall \boldsymbol{X} \in \Gamma^0 \\ \boldsymbol{u} &= \bar{\boldsymbol{u}} && \forall \boldsymbol{X} \in \partial\mathscr{B}^0_u. \end{aligned}$$

Here $\boldsymbol{P}$ is the first Piola-Kirchhoff stress, $\boldsymbol{N}$ is the unit normal vector to the boundary $\partial\mathscr{B}^0$ and $\boldsymbol{T}$ is the surface traction $\boldsymbol{t}$ with respect to an area element in the initial configuration.

Fig. 1 Schematic boundary value problem of cracked body in current configuration



2.2 Constitutive Equations for Finite Deformations Crystal Plasticity

Due to its strong relevance the modelling of crystal plasticity under finite deformations has a long tradition. Therefore extensive literature exists, and textbooks as well as overview articles cover this topic. Here just a short introduction to a simple crystal plasticity model is given. While the basic kinematic setup of continuum crystal plasticity theories is well established, see e.g. Asaro [1], there are still many experimental observable features as for example the deformation induced self organisation of dislocation patterns in dislocation cells, which are not yet fully understood. The constitutive model is embedded in the framework of elastic material isomorphism as e.g. motivated by Svendsen [28] or Bertram [6, 7].

Metals have a microstructure built of grains. A grain is modelled as ideal crystal with inherent oriented crystallographic planes. Plastic deformation is introduced as simple shear caused by sliding of several of this planes against each other keeping the crystalline structure unchanged. Because the crystalline symmetry is not changed the elastic behaviour also remains unchanged. If the elastic behaviour is assumed to be Hyperelastic, the second Piola-Kirchhoff stress $\boldsymbol{S}$ is defined as:

$$\boldsymbol{S} = 2\frac{\partial \psi_E}{\partial \boldsymbol{C}} \,, \tag{1}$$

with ψ representing the free energy density and $\boldsymbol{C}$ the right Cauchy-Green tensor. As commonly done the free energy density is assumed to decompose additively into an elastic and an isotropic hardening part and is of the form:

$$\psi := \psi(\boldsymbol{C}, \mathscr{P}, \gamma) = \psi_E(\boldsymbol{C}, \mathscr{P}) + \psi_{iso}(\gamma) \,. \tag{2}$$

Here $\{\mathscr{P}, \gamma\}$ are internal kinematic like variables, with $\mathscr{P}$ being the plastic transformation and γ being the accumulated plastic slip. If the elastic behaviour is not allowed to change during plastic deformation we obtain a material isomorphism and the dependence of the free energy on the right Cauchy-Green tensor and the plastic transformation has to be of the form: $\psi_E(\boldsymbol{C}, \mathscr{P}) = \psi_E(\mathscr{P}^T \boldsymbol{C} \mathscr{P})$. Equation (1) then reads:

$$\boldsymbol{S} = 2\frac{\partial \psi_E(\mathscr{P}^T \boldsymbol{C} \mathscr{P})}{\partial \boldsymbol{C}} = 2\mathscr{P} \underbrace{\frac{\partial \psi_E(\boldsymbol{C}_E)}{\partial \boldsymbol{C}_E}}_{=:S_E} \mathscr{P}^T \,, \tag{3}$$

with $\boldsymbol{C}_E := \mathscr{P}^T \boldsymbol{C} \mathscr{P}$.

As already mentioned, the plastic deformation is the combination of the slip of all active slip systems α. A slip system α is formed by the direction of the slip, modelled by the Euclidean vector $\boldsymbol{s}_\alpha$ and the plane α where the slip is acting defined by its normal vector $\boldsymbol{m}^\alpha$. Since they are orthogonal they can be understood as dual vectors of the α slip system:

$$\boldsymbol{s}_\alpha \cdot \boldsymbol{m}^\alpha = 0\ . \tag{4}$$

Here $\boldsymbol{s}_\alpha$ is a covariant and $\boldsymbol{m}^\alpha$ is a contravariant vector. With this definition and the rules for the pull-back of co- and contravariant vectors we can obtain the slip direction and the normal vector in the initial configuration by:

$$\begin{aligned} \boldsymbol{S}_\alpha &= \boldsymbol{F}^{-1}\boldsymbol{s}_\alpha \\ \boldsymbol{M}^\alpha &= \boldsymbol{F}^T\boldsymbol{m}^\alpha\ . \end{aligned} \tag{5}$$

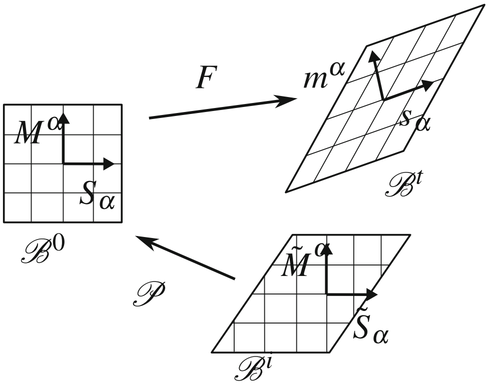
The introduction $\boldsymbol{s}_\alpha$ as covariant and $\boldsymbol{m}^\alpha$ as contravariant vector ensures their orthogonality under every deformation χ. Nevertheless if plastic deformation evolves, both, the Lagrangian vectors and the Euclidean vectors will also evolve. Therefore one can, as depicted in Fig. 2, pull back the slip system to a time independent isoclinic reference configuration:

$$\begin{aligned} \tilde{\boldsymbol{S}}_\alpha &= \mathscr{P}^{-1}\boldsymbol{S}_\alpha = \mathscr{P}^{-1}\boldsymbol{F}^{-1}\boldsymbol{s}_\alpha \\ \tilde{\boldsymbol{M}}^\alpha &= \mathscr{P}^T\boldsymbol{M}^\alpha = \mathscr{P}^T\boldsymbol{F}^T\boldsymbol{m}^\alpha\ . \end{aligned} \tag{6}$$

In order to decide if a slip system α is active the definition of the resolved shear stress τ^α is necessary. It is defined as the projection of the Cauchy stress $\boldsymbol{\sigma}$ onto the slip system α. With Eq. (6) and the application of the transformation rules for stresses the resolved shear stress is computed by:

$$\tau^\alpha := \boldsymbol{\sigma} : (\boldsymbol{m}^\alpha \otimes \boldsymbol{s}_\alpha) = \underbrace{\boldsymbol{C}_E\boldsymbol{S}_E}_{=:\boldsymbol{M}_E} : (\tilde{\boldsymbol{S}}_\alpha \otimes \tilde{\boldsymbol{M}}^\alpha)\ . \tag{7}$$

Fig. 2 Transformation of slip systems



The slip system α is active if the stress τ^α exceeds the yield limit $\tau_y^\alpha(\gamma)$. This leads to the definition of the yield surface as:

$$\Phi(\tau^\alpha, \tau_y^\alpha) := \max_\alpha (\Phi^\alpha) \leq 0 \qquad \text{with } \Phi^\alpha = \tau^\alpha - \tau_y^\alpha(\gamma) \ . \tag{8}$$

For simplicity we assume the same yield limit for all glide systems $\tau_y^\alpha(\gamma) = \tau_y(\gamma)$ and the isotropic hardening law:

$$\tau_y(\gamma) = \tau_{y0} + \frac{\partial \psi_{iso}}{\partial \gamma} \ . \tag{9}$$

To obtain a complete set of constitutive equations an evolution law for the internal variables $\{\mathscr{P}, \gamma\}$ needs to be defined. By the application of the principle of maximum dissipation we obtain the evolution equation for the plastic transformation:

$$\dot{\mathscr{P}} = -\mathscr{P} \underbrace{\sum_{\alpha=1}^{n} \dot{\gamma}_\alpha \tilde{\boldsymbol{S}}_\alpha \otimes \tilde{\boldsymbol{M}}^\alpha}_{=:\boldsymbol{\Lambda}} \ , \tag{10}$$

with $\dot{\gamma}$ being the slip rate of slip system α. In order to avoid an active set strategy we follow the method of visco-plastic regularisation similar to Steinmann and Stein in [27]. Therefore a constitutive equation for the slip rates is introduced:

$$\dot{\gamma}_\alpha = \dot{\gamma}_0 \frac{\tau^\alpha}{\tau_y} \left(\frac{< \tau^\alpha >}{\tau_y} \right)^{m-1} \ , \tag{11}$$

with the reference slip rate $\dot{\gamma}_0$, the sensitivity exponent m and $< x >$ being the Macauley operator. The total amount of slip can then be obtained by simple summation:

$$\dot{\gamma} = \sum_{\alpha=1}^{n_{sys}} \dot{\gamma}_\alpha \ . \tag{12}$$

The evolution equations (10) and (12) are integrated by a modified generalised midpoint rule.

2.3 Continuum Damage Mechanics Coupled to Crystal Plasticity

The modelling of failure and fracture in ductile materials is often modelled with the Gurson-Tvergaard-Needleman model (GTN) [29]. This model is micro mechanically motivated and captures the growth of a single pore in a perfect plastic matrix and additionally pore nucleation and coalescence. In [20] Needleman successfully studied the influence of pores around the crack tip and their impact to the fracturing process with the GTN model. Therefore the model is modified in this section to fit in the constitutive framework presented in Sect. 2.2.

In the GTN-model damage is characterised by the pore volume fraction f. The load carrying matrix material is reduced by this volume fraction and the yield strength is decreased. This is done by the introduction of a new yield surface Φ_f^α:

$$\Phi_f^\alpha := \left(\frac{\tau^\alpha}{\tau_y}\right)^2 - 1 + 2q_1 f^* \cosh\left(3q_2\frac{p}{\tau_y}\right) - (q_1 f^*)^2 \leq 0 \,. \tag{13}$$

The hydrostatic pressure p affects now the yield surface. While f^* is the effective void volume fraction, the parameters q_1 and q_2 are just model parameters and cannot be physically interpreted. The quantity f^* takes into account the void coalescence and nucleation if a material dependent threshold f_c is exceeded. It is determined by:

$$f^*(f) := \begin{cases} f, & \text{if } f \leq f_c \\ f_c + \frac{\frac{1}{q_1} - f_c}{f_F - f_c}(f - f_c) & \text{else .} \end{cases} \tag{14}$$

This relation considers the effect that the material strength decreases rapidly close to the final void volume fracture of failure f_F. In order to obtain the value of the internal variable f a new evolution equation has to be introduced. In [29] an additive split of f in a part f_g related to pore growth and f_n responsible for pore nucleation is proposed as:

$$f = f_g + f_n \,. \tag{15}$$

Originally the material is plastically incompressible, but by using the yield surface (13) a non deviatoric flow rule and thus a volumetric plastic deformation is introduced. The source of this volumetric deformation is the growth of pores and the respective evolution equation reads:

$$\dot{f}_g = (1 - f^*)\mathrm{Tr}(\mathscr{P}^{-1}\dot{\mathscr{P}}) \,. \tag{16}$$

The pore nucleation part of the evolution equation follows a more phenomenological argumentation. Since it is a statistical process a standard distribution is assumed as:

$$\dot{f}_n = \frac{A_n}{\sqrt{2\pi\sigma_n}} \exp\left(-\frac{(\gamma - \gamma_n)^2}{2\sigma_n}\right)\dot{\gamma} \,. \tag{17}$$

This constitutive law for void nucleation is strain controlled by the mean strain for pore nucleation γ_n and has the variance σ_n. A further extension of the presented model is the introduction of a stress controlled constitutive law for the pore nucleation as suggested in [20]. To complete the set of constitutive equation the free energy function ψ is specified as:

$$\begin{aligned}\psi(\boldsymbol{C}_E, \gamma) =& \frac{1}{2}\mu(\mathrm{Tr}\boldsymbol{C}_{Eiso} - 3) + \frac{\kappa}{8}(\ln(\det(\boldsymbol{C}_E)))^2 \\ &+ \frac{1}{2}H\gamma^2 + h1\left(\frac{\exp(-h_2\gamma)}{h_2} + \gamma\right)\tau_{y0}.\end{aligned} \tag{18}$$

2.4 Non-Local Extension of the GTN Model

If the boundary value problem BVP 1 with the constitutive equations of Sects. 2.2 and 2.3 is solved with the finite element method the problem of mesh dependent localization occurs. This was observed quite early by Pietruszczak and Mróz [24]. This mesh dependency typically results in a strong dependence of the force displacement curve on the element size. Instead of convergence to a certain material response, softening materials tend to diverge with decreasing element size. Another example is the formation of localisation bands. Here the width of the localised deformation is close to the element size. So if some material microstructure constrains the width of such localisation this has to be taken into account explicitly.

Several attempts have been made to circumvent the above described behaviour. Here we focus on the gradient dependent theories which are associated with the names micromorphic continua [10] or gradient enhanced damage models [23] and others. All these models rely on the introduction of a new field variable ϕ and its gradient $\nabla\phi$ and the coupling of the local evolution of an internal variable to these quantities. These theories only differ in their motivation and derivation, but end up with similar additional balance equation. Reusch, Svendsen and Klingbeil proposed in [26] such a theory relating the evolution of the volume void fraction $\dot{f}$ of the GTN model to its non-local equivalent $\dot{\phi}$. This new field variable has to fulfil the equation:

$$\begin{aligned}\dot{\phi} - \ell^2\Delta\phi &= \dot{f} \qquad \forall \boldsymbol{X} \in \mathscr{B}^0 \\ \nabla\phi \cdot \boldsymbol{N} &= 0 \qquad \forall \boldsymbol{X} \in \partial\mathscr{B}^0,\end{aligned} \tag{19}$$

with homogeneous Neumann boundary condition. Here the local quantity $\dot{f}$ forms the right hand side of the equation and couples the pure displacement based boundary value problem BVP 1 with the non-local part. The material parameter ℓ is related to the size of microscopic material feature and can be determined from the width of a real measured localisation band. In the example Sect. 5.1 a simple test problem is studied for different mesh resolutions.

2.5 *Weak Form of the Resulting Boundary Value Problem*

To summarize the equation introduced so far we will present the weak form of the boundary value problem BVP 1 and Eq. (19). This weak form is given in the initial configuration. By the standard procedure of multiplying this both equations with some arbitrary functions $\{\delta \boldsymbol{u}, \delta\phi\}$, integrating over the entire domain and applying the divergence theorems one ends up with:

BVP 2 *Find* $\boldsymbol{u} \in \mathscr{U}$ *and* $\phi \in \mathscr{H}^1(\mathscr{B})$ *such that for all* $\delta\boldsymbol{u} \in \mathscr{V}$ *and* $\delta\phi \in \mathscr{H}_0^1$*:*

$$\mathscr{G}_u(\boldsymbol{u}, \delta\boldsymbol{u}) = \int_{\mathscr{B}} \delta\boldsymbol{F} : \boldsymbol{P}\, dV - \int_{\partial\mathscr{B}_\sigma} \delta\boldsymbol{u}\tilde{\boldsymbol{T}} dA = 0$$

$$\mathscr{G}_\phi(\phi, \delta\phi) = \int_{\mathscr{B}} \delta\phi(\dot{f} - \dot{\phi}) - \ell^2 \nabla\delta\phi \cdot \nabla\phi dV = 0$$

subjected to the boundary condition:

$$\begin{aligned} \boldsymbol{PN} &= \bar{\boldsymbol{T}} && \forall \boldsymbol{X} \in \partial\mathscr{B}_\sigma^0 \\ \boldsymbol{PN} &= \boldsymbol{0} && \forall \boldsymbol{X} \in \Gamma^0 \\ \boldsymbol{u} &= \bar{\boldsymbol{u}} && \forall \boldsymbol{X} \in \partial\mathscr{B}_u^0 \\ \nabla\phi \cdot \boldsymbol{N} &= 0 && \forall \boldsymbol{X} \in \partial\mathscr{B}^0 \end{aligned}$$

is fulfilled. Where the vector spaces $\mathscr{U}$ *and* $\mathscr{V}$ *are defined as:*

$$\mathscr{U} = \left\{\boldsymbol{u} | \boldsymbol{u} \in \left(\mathscr{H}^1(\mathscr{B})\right)^3\right\} \qquad \mathscr{V} = \left\{\delta\boldsymbol{u} | \delta\boldsymbol{u} \in \left(\mathscr{H}^1(\mathscr{B})\right)^3, \delta\boldsymbol{u} = \boldsymbol{0} \text{ on } \partial\mathscr{B}_u^0\right\} .$$

3 Discretisation of the Boundary Value Problem with the XFEM

In this section the discretisation of the weak form of the boundary value problem BVP 2 with the eXtended Finite Element method (XFEM) [4] is described. The XFEM is used because the displacement field is discontinuous across the crack surface Γ and this method is capable to handle this accurately with just a minor additional computational expense.Within the XFEM additional degrees of freedom are introduced to the nodes in the vicinity of the crack. This enriches the approximation space of the displacement by some knowledge of the characteristic features of the solution. Along the crack surface the nodes become jump enriched because of the presence of a displacement jump in the solution. In linear elastic fracture mechanics, at the crack front the nodes are usually enriched by the basis functions of the asymptotic solution. In Fig. 3 the enrichment scheme for a 2D model

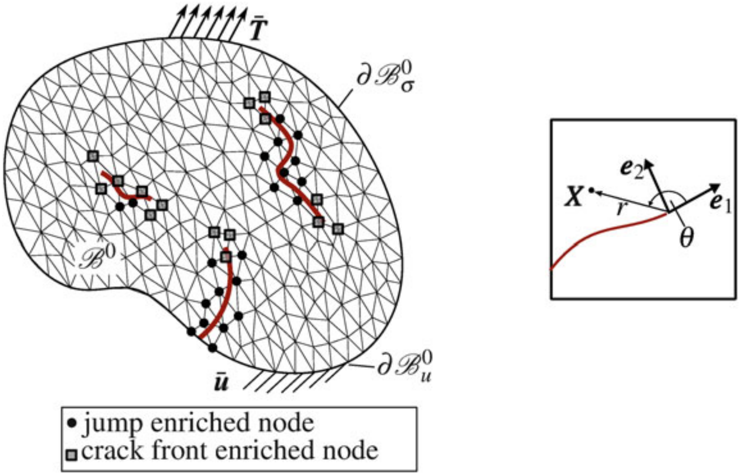


Fig. 3 Discretised boundary value problem (*left*) and local coordinate system at the crack tip (*right*)

problem is given. It is obvious that the total number of degrees of freedom is just increased moderately.

The main advantage of the XFEM is the fact, that in case of crack propagation remeshing is not necessary due to the nearly mesh independent description of the crack geometry and the displacement discontinuity. Usually cracks are implicitly modelled by two mutually perpendicular level set functions φ_1 and φ_2. While the level set φ_1 represents the signed distance to the crack surface, the second level set φ_2 describes the signed distance to the crack front. With this convention the jump enrichment function is defined as:

$$H(\boldsymbol{X}) := \begin{cases} 1, & \text{for } \varphi_1 \geq 0 \\ -1, & \text{for } \varphi_1 < 0\,. \end{cases} \tag{20}$$

Elguedj et al. [9] derived an enrichment basis for the crack front field for the case of small deformation elasto-plasticity with power law hardening based on the well known HRR-field. Since there is no analytic solution for the near tip field in case of finite deformations crystal plasticity we follow Legrain et al. [16] and Loehnert et al. [17] and just employ one enrichment function f_1. The enrichment function reads:

$$f_1(r,\theta) = r^n \sin\left(\frac{\theta}{2}\right)\,, \tag{21}$$

where r and θ describes the crack geometry as seen in Fig. 3 on the right hand side. The same holds for the three dimensional case where at every point of the crack front such a local coordinate system is placed. With n chosen to be $n \geq 1$ the derived strain and stress fields do not have a singularity if $r \to 0$. With this introduction the approximated displacement field $\boldsymbol{u}^h(\boldsymbol{X})$ is computed via:

$$\begin{aligned} \boldsymbol{u}^h(\boldsymbol{X}) &= \sum_{I=1}^{n} N_I(\boldsymbol{X}) \left[\boldsymbol{u}_I + H(\boldsymbol{X})\boldsymbol{a}_I + \sum_{j=1}^{m} f_j(\boldsymbol{X})\boldsymbol{b}_{jI} \right] \\ &= \sum_{I=1}^{n} \left[N_I \; N_I H \; N_I f_1 \; \ldots \; N_I f_m \right] \begin{bmatrix} \boldsymbol{u}_I \\ \boldsymbol{a}_I \\ \boldsymbol{b}_{1I} \\ \ldots \\ \boldsymbol{b}_{mI} \end{bmatrix} = \sum_{I=1}^{n} \mathcal{N}_I^T \mathbf{u}_I \,. \end{aligned} \tag{22}$$

Since the non-local void volume fraction ϕ can develop independently on both sides of the crack and is assumed to have a peak value at the crack front similar enrichment functions are used:

$$\phi^h(\boldsymbol{X}) = \sum_{I=1}^{n} N_I(\boldsymbol{X}) \left[\phi_I + H(\boldsymbol{X})\phi_{aI} + \sum_{j=1}^{m} f_j^{\phi}(\boldsymbol{X})\phi_{bjI} \right] = \mathcal{N}_I^T \boldsymbol{\phi}_I \,. \tag{23}$$

By employing a Bubnov-Galerkin scheme the same Ansatz is used for the test functions $\delta\mathbf{u}^h$ and $\delta\phi^h$. The discretised versions of the weak forms of BVP 2 are then computed by:

$$\begin{aligned} \mathcal{G}_u(\mathbf{u}^h, \delta\mathbf{u}^h) &= = \bigcup_{e=1}^{n_e} \sum_{I=1}^{N} \delta\mathbf{u}_I \underbrace{\int_{\mathcal{B}_e} \nabla\mathcal{N}_I^T \boldsymbol{P} \, dV_e}_{\boldsymbol{P}_{u_I}} - \sum_{I=1}^{N} \delta\mathbf{u}_I \underbrace{\int_{\partial\mathcal{B}_e} \mathcal{N}_I \tilde{\boldsymbol{T}} \, dA_e}_{F_{\text{ext}_I}} \\ \mathcal{G}_\phi(\boldsymbol{\phi}^h, \delta\boldsymbol{\phi}^h) &= = \bigcup_{e=1}^{n_e} \sum_{I=1}^{N} \delta\boldsymbol{\phi}_I \underbrace{\int_{\mathcal{B}_e} \mathcal{N}_I^T (\dot{f} - \dot{\phi}) - \ell^2 \nabla\mathcal{N}_I^T \cdot \nabla\phi dV}_{P_{\phi_I}} \,. \end{aligned} \tag{24}$$

The nonlinear set of equations is treated by a standard Newton-Raphson procedure. In each iteration the linear system:

$$\begin{bmatrix} \boldsymbol{F}_{\text{ext}} - \boldsymbol{P}_u \\ \boldsymbol{P}_\phi \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_{uu} & \boldsymbol{K}_{u\phi} \\ \boldsymbol{K}_{\phi u} & \boldsymbol{K}_{\phi\phi} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\boldsymbol{\phi} \end{bmatrix} \tag{25}$$

is solved where $\boldsymbol{K}_{ij}$ is the consistent linearisation of the weak form i with respect to the degree of freedom j with $i, j \in \{\mathbf{u}, \boldsymbol{\phi}\}$.

Because of the complexity of real structures we discretise the domain by 10 node tetrahedral elements with quadratic shape functions. The integration is performed

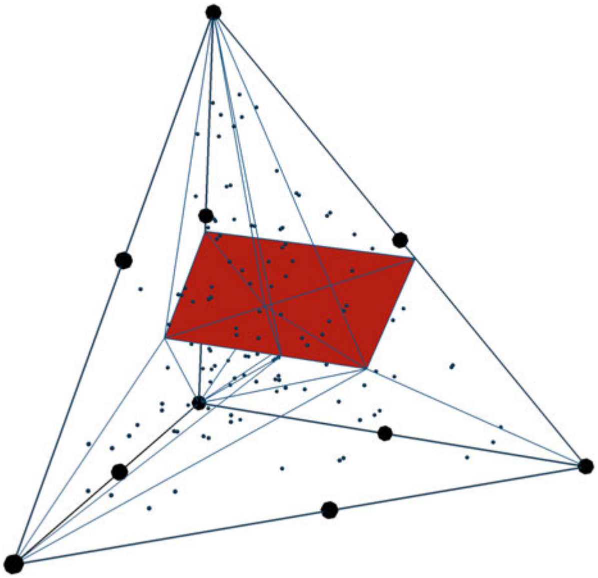


Fig. 4 Integration scheme for an inclined crack (*red*) within a single 10 node tetrahedral element

by a subcell strategy where the element is split along the discontinuity. A standard Gauss integration is then used in each subcell as can be seen in Fig. 4.

4 Quasistatic Crack Propagation with the XFEM

The task of propagating cracks is here divided into three steps. After a short motivation of the crack propagation criterion the crack tracking algorithm is presented. The section is concluded by embedding the crack propagation algorithm into a standard load stepping scheme of finite element software.

4.1 Crack Propagation Criterion

Ductile fracture occurs if the material strength is exceeded. The material strength is measured for a defect free mono crystal and is much higher than the strength of real materials. For brittle materials in linear elastic fracture mechanics the concept of a critical energy release rate $\mathscr{G}_c$ is used. This concept goes back to the works of Griffith [13], where $\mathscr{G}_c$ is the required energy to form a new Surface ΔA. This criterion is based on the assumption of linear elasticity and small deformations and cannot be transferred to the current framework easily. Therefore we treat crack propagation as the coalescence of voids close to the crack tip. Instead of using an energy based threshold we propose to use a critical void volume fraction f_{prop} to propagate the crack. In [20] and [3] Needleman et al. also used the void volume fraction to predict the crack path. But their approach suffers from mesh depended crack paths and numerical difficulties when $f \to f_F$ because the problem becomes ill posed. With the here presented strategy we are able to treat cracks sharply and nearly mesh independent. Following the damage concept the crack will propagate in the direction where voids grow fastest. This direction can be identified to be proportional to $\nabla\dot{\phi}$. In the following sections an efficient algorithm is presented to propagate the crack.

4.2 Reconstruction of the Fracture Surface

The presented algorithms do not depend on the choice of crack propagation direction, so this direction is substituted by the vector field $\boldsymbol{v}(\boldsymbol{X})$. In principle there are two ways of reconstructing the crack surface after determining the crack velocity vectors $\boldsymbol{v}$. The first is an explicit reconstruction of the surface by discretising the crack front line and computing the velocity vectors $\boldsymbol{v}$ at each point. By joining the tips of the vectors the new surface is determined. This approach was successfully implemented in a three dimensional context by Fries et al. [11] or by Holl et al. [14]. However, for inelastic materials this technique oscillatory behaviour of the crack front line geometry and after a few propagation steps leads to this results in unphysical crack geometries.

The second approach is to implicitly track the crack front, by either using the level set approach as done by Moës and Gravouil in Moës et al. [19] and Gravouil et al. [12] or by the global crack tracking algorithm suggested by Oliver and Huespe in [22]. We will present both approaches and study their accuracy in the following sections.

4.2.1 Level Set Update by Solving a Hamilton-Jacobi Equation

As already described in Sect. 3 the cracks are represented by two level set functions φ_1, φ_2. The evolution of such fields follows the Hamilton-Jacobi equation:

$$\begin{aligned} \tfrac{\partial\varphi_i}{\partial t} + \bar{\boldsymbol{v}}_i \cdot \nabla\varphi_i &= \bar{g} \quad \forall \boldsymbol{X} \in \mathscr{B}^0 \\ \varphi_i \qquad &= \bar{\varphi}_i \ \forall \boldsymbol{X} \in \partial\mathscr{B}^0_{\varphi} \ , \end{aligned} \tag{26}$$

where $\bar{\boldsymbol{v}}_i$ is the velocity of the i-th level set. Equation (26) holds for different purposes, just by replacing the velocity $\bar{\boldsymbol{v}}_i$ or the right hand side $\bar{g}$ according to Table 1. In Barths and Sethians contribution [2] they proposed an explicit method based on a Petrov-Galerkin formulation for the solution of Eq. (26). The proposed algorithm is designed for simplex elements. Since we want to deal with higher order

Table 1 Velocities and right hand sides for different purposes

Purpose	$\bar{\boldsymbol{v}}$	$\bar{g}$
φ_1 propagation	$\boldsymbol{v}_{\varphi_1}$	0
φ_2 propagation	$\boldsymbol{v}_{\varphi_2}$	0
Reinitialisation	$\frac{\mathrm{sign}(\varphi_1)^0 \nabla\varphi 1}{\|\|\nabla\varphi_1\|\|}$	$\mathrm{sign}(\varphi_1)^0$
Reorthogonalisation	$\frac{\mathrm{sign}(\varphi_1) \nabla\varphi 1}{\|\|\nabla\varphi_1\|\|}$	0

elements we derived an implicit algorithm by applying the Galerkin Least Square (GLS) formulation of Hughes et al. [15] for advection dominated problems. The weighted residual corresponding to Eq. (26) reads:

$$\mathscr{G}_{\varphi}^{G}(\varphi, \delta\varphi) = \int_{\mathscr{B}^0} \delta\varphi \left(\frac{\partial\varphi}{\partial t}\right) + \delta\varphi \bar{\boldsymbol{v}} \cdot \nabla\varphi - \delta\varphi g dV = 0\ . \tag{27}$$

To this weighted residual the least square approximation $\delta\Pi_{\varphi}^{LS}$ of the residual R is added:

$$\begin{aligned} R &:= \sum_{I=1}^{N} N_I \dot{\varphi}_I + \sum_{I=1}^{N} \bar{\boldsymbol{v}} \cdot \nabla N_I \varphi_I - \bar{g} \\ \mathscr{G}_{\varphi}^{LS}(\varphi, \delta\varphi) &:= \bigcup_{e}^{n_e} \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}\varphi_I} \int_{\mathscr{B}^0} \tau R^2 \mathrm{d}V \\ &= \bigcup_{e}^{n_e} \int_{\mathscr{B}^0} \sum_{I=1}^{N} \left(N_I \frac{\partial\dot{\varphi}_I}{\partial\varphi_I} + \bar{\boldsymbol{v}} \cdot \nabla N_I\right) \tau \sum_{J=1}^{N} \left(N_J \dot{\varphi}_J + \bar{\boldsymbol{v}} \cdot \nabla N_J \varphi_J - \bar{g}\right) \mathrm{d}V\ . \end{aligned} \tag{28}$$

The Parameter τ in Eq. (28) is a stabilisation parameter which depends on the element size h and whose value is chosen after [2]:

$$\tau = \left(\left[\frac{2}{\Delta t}\right]^2 + \left[\frac{2||\bar{\boldsymbol{v}}||}{h}\right]^2\right)^{-\frac{1}{2}}\ . \tag{29}$$

The resulting weighted residual is in general a non linear equation of the form:

$$\mathscr{G}_{\varphi}^{GLS} = \mathscr{G}_{\varphi}^{G} + \mathscr{G}_{\varphi}^{LS}\ , \tag{30}$$

where just the Galerkin part is consistently linearised in a Newton-Raphson scheme. Depending on the used time discretisation technique this equation can be solved explicitly or implicitly. Because of its higher order accuracy and its easy implementation we chose a Crank-Nicolson discretisation.

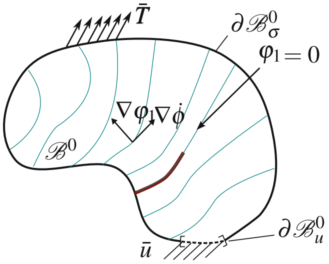


Fig. 5 BVP for global crack tracking

4.2.2 Global Crack Tracking Algorithm

The crack surface is described implicitly by the level set value φ_1. Oliver and Huespe [22] postulate the existence of two linearly independent vector fields $\boldsymbol{T}$ and $\boldsymbol{S}$ which are perpendicular to the normal $\boldsymbol{n} = \frac{\nabla\varphi_1}{||\nabla\varphi_1||}$ of the searched surface. This situation is depicted in Fig. 5 for the two dimensional case. The problem is to find some φ_1 so that:

$$\begin{aligned} \boldsymbol{T}\cdot\nabla\varphi_1 &= 0 \\ \boldsymbol{S}\cdot\nabla\varphi_1 &= 0 \end{aligned} \tag{31}$$

is fulfilled. Multiplying both sides of the equation with $\boldsymbol{T}$ or $\boldsymbol{S}$ does not alter the problem.

$$\begin{aligned} (\boldsymbol{T}\otimes\boldsymbol{T})\cdot\nabla\varphi_1 &= \boldsymbol{0} \\ (\boldsymbol{S}\otimes\boldsymbol{S})\cdot\nabla\varphi_1 &= \boldsymbol{0} \end{aligned} \tag{32}$$

We now add up these two equation and take the divergence. This will lead to a boundary value problem with homogeneous Neumann boundary conditions:

BVP 3 *Find φ_1 so that:*

$$\begin{aligned} \mathrm{Div}\,(-\left[\boldsymbol{T}\otimes\boldsymbol{T}+\boldsymbol{S}\otimes\boldsymbol{S}\right]\cdot\nabla\varphi_1) = \mathrm{Div}\,(-\mathbb{K}\cdot\nabla\varphi_1) &= 0 \qquad && \forall\boldsymbol{X}\in\mathscr{B}^0 \\ -\mathbb{K}\cdot\nabla\varphi_1 &= \boldsymbol{q} && \forall\boldsymbol{X}\in\mathscr{B}^0 \\ \boldsymbol{q}\cdot\boldsymbol{N} &= 0 && \forall\boldsymbol{X}\in\partial\mathscr{B}^0_q \\ \varphi &= \bar{\varphi} && \forall\boldsymbol{X}\in\partial\mathscr{B}^0_\varphi \end{aligned} \tag{33}$$

is fulfilled. The Dirichlet bound $\partial\mathscr{B}^0_\varphi$ is chosen to be the old crack path with negative second level set.

This simple heat equation like linear boundary value problem can be written in its discretised weak form:

$$\mathscr{G}_{\varphi}^{GCT}(\varphi, \delta\varphi) = \bigcup_{e}^{n_e} \delta\varphi_I \sum_{I=1}^{N} \sum_{J=1}^{N} \int_{\mathscr{B}^0} \nabla N_I \mathbb{K} \nabla N_J dV \varphi_J = 0\ . \tag{34}$$

It should be remarked that the tensor $\mathbb{K}$ shows up a rank deficiency of 1. This can lead to an ill conditioned equation system. Therefore a small multiple ε of the zero eigen space is added to $\mathbb{K}$. The influence of this stabilisation parameter on the solution is later studied for an example problem. For $\boldsymbol{T}$ and $\boldsymbol{S}$ an adequate choice needs to be made. Since the crack propagation vector $\boldsymbol{v}$ is tangential to the crack surface, $\boldsymbol{T}$ is chosen to be equal to $\boldsymbol{v}$. $\boldsymbol{S}$ is assumed to be the crossproduct between the eigenvector of the smallest principle strain $\boldsymbol{E}_1$ with $\boldsymbol{v}$:

$$\begin{aligned} \boldsymbol{T} &= \boldsymbol{v} \\ \boldsymbol{S} &= \boldsymbol{v} \times \boldsymbol{E}_1\ . \end{aligned} \tag{35}$$

For the second level set φ_2 this update technique is not possible because the need for an adequate Dirichlet boundary condition. Therefore the Petrov-Galerkin scheme is used to update the second level set.

4.2.3 Accuracy of the Two Methods

In this section we study the accuracy of the describes models for propagating the crack surface in a simple example similar to the study of Duflot [8]. In a block of size 1 × 1 × 0.25 a horizontal crack is placed. The block is discretised by 9600 tetrahedral elements with quadratic shape functions and the crack velocity field $\boldsymbol{v}$ is prescribed to enforce a crack path inclined by the angle α. In Fig. 6 on the left hand side the setup is displayed.

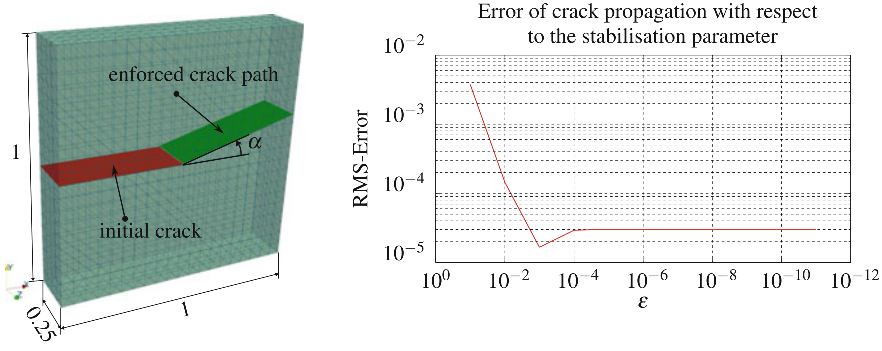


Fig. 6 Setup of the test problem (*left*) and Deviation of prescribed crack path with respect to stabilisation parameter for $\alpha = 30°$ (*right*)

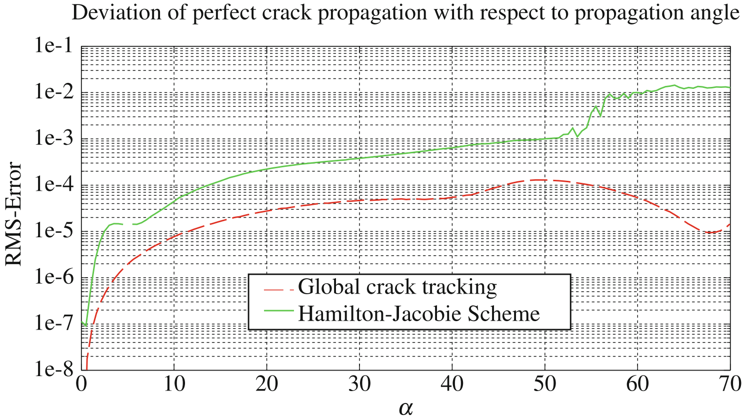


Fig. 7 Deviation of prescribed crack path with respect to the propagation angle

In the first test the stabilisation parameter ε of the global crack tracking algorithm is changed and the root mean square error:

$$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\varphi}_i - \varphi)^2} \tag{36}$$

of the resulting crack with respect to the expected crack path is evaluated. Here $\hat{\varphi}_i$ is the signed distance of the analysed method which is compared with the exact one φ. As can be seen in Fig. 6 on the right hand side the error becomes minimal for a stabilisation parameter of the order $\varepsilon \approx 10^{-3}$. For smaller parameters the deviation remains nearly constant. In the ongoing examples we will therefore use $\varepsilon = 10^{-3}$.

In a second example the angle α is altered in the range $0° \leq \alpha \leq 70°$. The level set update is performed by using the implicit Petrov-Galerkin scheme and we compare these results with those obtained using the global crack tracking algorithm. In Fig. 7 the root mean square error of both methods with respect to the propagation angle is displayed. The global tracking algorithm performs better in the entire range. Over a broad range $0° \leq \alpha \lesssim 53°$ the error of this algorithm is one order of magnitude smaller than the error of the Petrov-Galerkin scheme. For larger angles the discrepancy is even up to two orders of magnitude. We therefore recommend to use the global tracking algorithm for updating the first level set.

4.3 *Crack Propagation Algorithm*

In this section a simple displacement controlled algorithm for quasi static crack propagation is presented. After the level set update the crack propagation algorithm continues in updating the enrichment scheme according to the new level sets. By

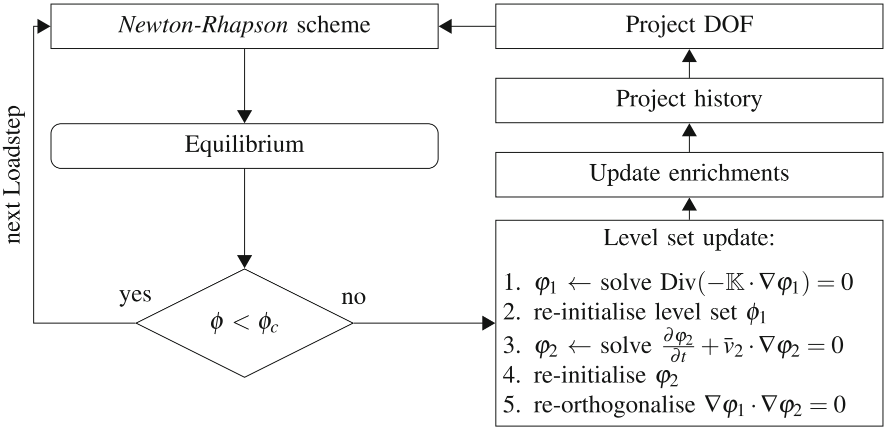


Fig. 8 Scheme for propagating cracks

doing this the mesh will not be changed but the nodal degrees of freedom close to the crack front will alter. Also the integration point scheme will change to take the new crack geometry into account. It is necessary to project the displacement field described with the old enrichment pattern onto the description using the new enrichment pattern. Similarly it is also necessary to project the history variables from the old integration points onto the new integration points in those parts of the domain where the position of the integration points change. For both tasks a linear least square problem is solved. Since now neither the yield condition nor the mechanical equilibrium is fulfilled an equilibrium iteration is performed. After equilibrium is achieved the test for the critical void volume fraction is performed again. This procedure is repeated as long as the critical volume fraction is exceeded. A schematic representation of the algorithm is given in Fig. 8.

Remark 1 It has to be remarked that the update of the history variables and the projection of the degrees of freedom is still the weak part the proposed algorithm. Often the trial values for constitutive material routine are to far away from the solution after a crack propagation step. To solve this issue a more advanced projection technique for the displacement field and the history variables and a new strategy for integrating the constitutive equations is under development. For simple examples and small crack propagation increments the algorithm is stable.

5 Numerical Examples

In this section three examples for the described methods are presented. After a verification of the non local GTN formulation two simple crack propagation examples are considered.

5.1 Verification of the Non-Local Damage Formulation

In Fig. 9 a polycrystalline block with given dimension is shown. The block is a composition of Voronoi cells generated with the software Neper [25] and each cell is representing a grain of the material. The orientation of the grains is distributed randomly and displayed as different colour in Fig. 9. The crystal lattice is body centred cubic with 12 slip systems according to Nemat-Nasser et al. [21] and the material data can be found in Table 2 which correspond to a DC04 steel. A displacement controlled tensile test without constraining the lateral stretch is performed and the external force and displacement is recorded. In Fig. 10 the non-local void volume fraction is shown. In the undeformed configuration on the left

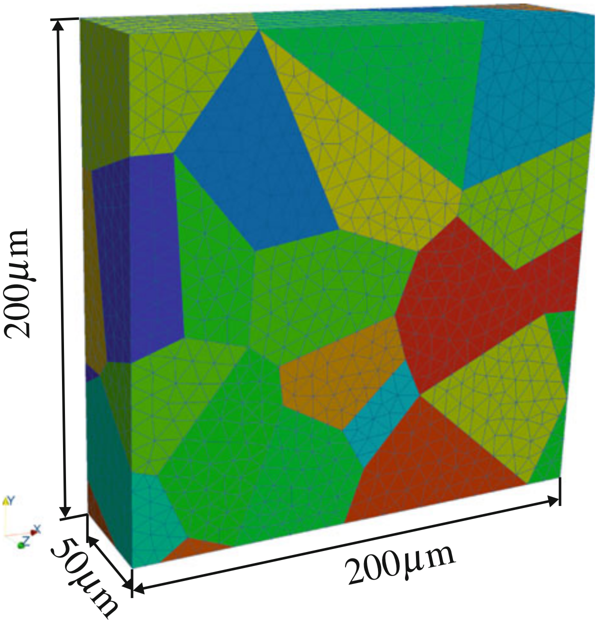


Fig. 9 Setup for mesh dependency test

Table 2 Material and model parameter

Parameter	Value	Parameter	Value
μ	80,000 MPa	q_1	1.5
κ	160,000 MPa	q_2	1
τ_{y0}	80 MPa	ℓ	0.03 mm
$h_1 = h_2$	1	f_c	0.07
H	0	σ_n	0.2
m	3	A_n	0.5
$\dot{\gamma}_0$	0.005		
f_F	0.3		

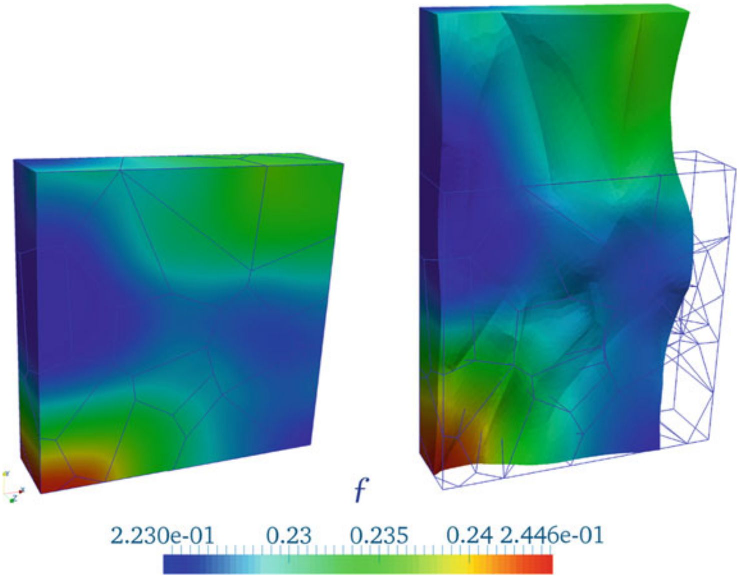


Fig. 10 Void volume fraction in undeformed state (*left*) and in deformed state (*right*)

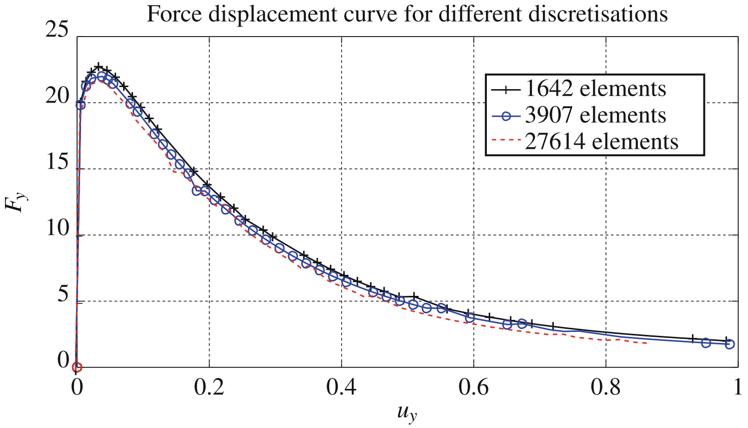


Fig. 11 Force displacement curve of a tensile test with different discretisations

hand side one can see that there are grains which promote the damage evolution for the given load (red and green) and a band of grains (blue) which has a higher remaining strength. Furthermore the distribution of the void volume fraction is very smooth which reflects its non-local nature. This test is repeated for 3 different discretisations where the element size is decreasing. The results are displayed in Fig. 11 and show a convergence behaviour to a certain load deflection path. The presented formulation is clearly mesh independent. The validation of the model is ongoing research where a real material testing will be performed for comparison.

5.2 *Crack Propagation in Crystals*

The next test example considers a homogeneous specimen with a horizontal crack and dimensions shown in Fig. 12 on the left hand side. The block is treated as single crystal and a vertical displacement is prescribed on the top surface. In order to not constrain the lateral contraction symmetric boundary conditions are applied. The example is uniformly discretised with 2592 tetrahedral elements with quadratic shape functions and the material and model parameters can be found in Table 2. For this setup the crack has to propagate horizontally which is in fact the case as can be seen in Fig. 12 on the right hand side.

By crossing a grain boundary the direction of crack propagation should change according to the lattice orientation. Since the crack is driven by the volume void distribution and the strain state the presented model should be able to reflect this behaviour implicitly. To study this effect the block of the previous examples is now composed of two grains with different orientation as can be seen in Fig. 13. The discretisation and the boundary conditions are the same as before. On the right hand side of Fig. 13 the simulated crack path is shown. Close to the material interface the direction of the crack propagation changes. The crack grows into the second grain and proceeds its propagation with a different direction.

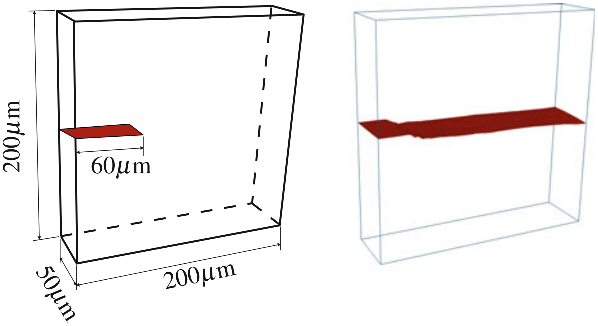


Fig. 12 Setup for crystalline crack growth (*left*) and simulated crack path for tensile test (*right*)

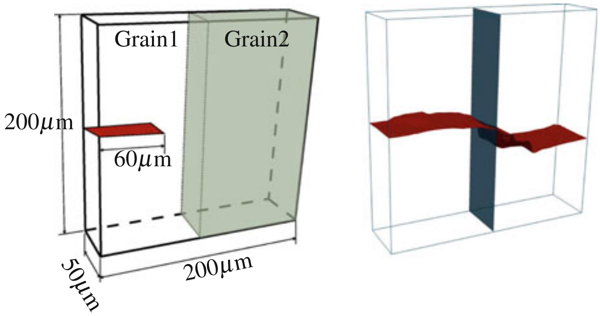


Fig. 13 Setup for inter crystalline crack growth (*left*) and simulated crack path for tensile test (*right*)

6 Conclusions

In this work the entire process of formulating the boundary value problem, deriving the constitutive equation and the numerical treatment of fracture in polycrystals is derived. Therefore a non-local extension of the Gurson-Tvergaard-Needleman model in finite deformations crystal plasticity is introduced. In mesh dependence study we show that the presented model does not suffer from mesh dependent localisation.

To represent cracks sharply a transition from damage to fracture is described. The cracks are modelled with the extended finite element method and two update techniques for the level sets are investigated. Here the global crack tracking scheme is recommended for updating the first level set because of its simplicity and accuracy. For re-initialisation purposes and for updating the second level set an implicit Petrov-Galerkin scheme is derived. Finally a displacement controlled crack propagation algorithm is outlined and some open questions are remarked. The problem of mapping history values after a crack propagation step and the corresponding issue of updating the constitutive law is addressed. In three examples the features of the presented model are shown.

Acknowledgements This work is supported by the German Research Foundation (DFG) under the Transregional Collaborative Research Center SFB/TR73: “Manufacturing of Complex Functional Components with Variants by Using a New Sheet Metal Forming Process—Sheet-Bulk Metal Forming”.

References

1. Asaro, R.J.: Crystal plasticity. J. Appl. Mech. **50**, 921–934 (1983)
2. Barth, T.J., Sethian, J.A.: Numerical schemes for the Hamilton-Jacobi and level set equations on triangulated domains. J. Comput. Phys. **145**, 1–40 (1998)

3. Becker, R., Needleman, A., Suresh, S., Tvergaard, V., Vasudevan, A.: An analysis of ductile failure by grain boundary void growth. Acta Metall. **37**, 99–120 (1989)
4. Belytschko, T., Black, T.: Elastic crack growth in finite elements with minimal remeshing. Int. J. Numer. Methods Eng. **45**, 601–620 (1999)
5. Belytschko, T., Moës, N., Usui, S., Parimi, C.: Arbitrary discontinuities in finite elements. Int. J. Numer. Methods Eng. **50**, 993–1013 (2001)
6. Bertram, A.: An alternative approach to finite plasticity based on material isomorphisms. Int. J. Plast. **15**, 353–374 (1999)
7. Bertram, A.: Elasticity and Plasticity of Large Deformations. Springer, Berlin (2008)
8. Duflot, M.: A study of the representation of cracks with level sets. Int. J. Numer. Methods Eng. **70**, 1261–1302 (2007)
9. Elguedj, T., Gravouil, A., Combescure, A.: Appropriate extended functions for X-FEM simulation of plastic fracture mechanics. Comput. Methods Appl. Mech. Eng. **195**, 501–515 (2006)
10. Eringen, A.C., Suhubi, E.: Nonlinear theory of simple micro-elastic solids i. Int. J. Eng. Sci. **2**, 189–203 (1964)
11. Fries, T.-P., Baydoun, M.: Crack propagation with the extended finite element method and a hybrid explicit implicit crack description. Int. J. Numer. Methods Eng. **89**, 1527–1558 (2012)
12. Gravouil, A., Moës, N., Belytschko, T.: Non-planar 3d crack growth by the extended finite element and level sets part ii: level set update. Int. J. Numer. Methods Eng. **53**, 2569–2586 (2002)
13. Griffith, A.A.: The phenomena of rupture and flow in solids. Philos. Trans. R. Soc. Lond. Ser. A, Contain. Pap. Math. Phys. Character **221**, 163–198 (1921)
14. Holl, M., Rogge, T., Loehnert, S., Wriggers, P., Rolfes, R.: 3d multiscale crack propagation using the xfem applied to a gas turbine blade. Comput. Mech. **53**, 173–188 (2014)
15. Hughes, T.J., Franca, L.P., Hulbert, G.M.: A new finite element formulation for computational fluid dynamics. VIII: the Galerkin/least-squares method for advective-diffusive equations. Comput. Methods Appl. Mech. Eng. **73**, 173–189 (1989)
16. Legrain, G., Moës, N., Verron, E.: Stress analysis around crack tips in finite strain problems using the extended finite element method. Int. J. Numer. Methods Eng. **63**, 290–314 (2005)
17. Loehnert, S., Mueller-Hoeppe, D., Wriggers, P.: 3d corrected xfem approach and extension to finite deformation theory. Int. J. Numer. Methods Eng. **86**, 431–452 (2011)
18. Moës, N., Dolbow, J., Belytschko, T.: A finite element method for crack growth without remeshing. Int. J. Numer. Methods Eng. **46**, 131–150 (1999)
19. Moës, N., Gravouil, A., Belytschko, T.: Non-planar 3d crack growth by the extended finite element and level sets part I: mechanical model. Int. J. Numer. Methods Eng. **53**, 2549–2568 (2002)
20. Needleman, A., Tvergaard, V.: An analysis of ductile rupture modes at a crack tip. J. Mech. Phys. Solids **35**, 151–183 (1987)
21. Nemat-Nasser, S., Okinaka, T., Ni, L.: A physically-based constitutive model for BCC crystals with application to polycrystalline tantalum. J. Mech. Phys. Solids **46**, 1009–1038 (1998)
22. Oliver, J., Huespe, A.: Continuum approach to material failure in strong discontinuity settings. Comput. Methods Appl. Mech. Eng. **193**, 3195–3220 (2004); Computational failure mechanics
23. Peerlings, R.H.J., de Borst, R., Brekelmans, W.A.M., de Vree, J.H.P.: Gradient enhanced damage for quasi-brittle materials. Int. J. Numer. Methods Eng. **39**, 3391–3403 (1996)
24. Pietruszczak, S., Mróz, Z.: Finite element analysis of deformation of strain-softening materials. Int. J. Numer. Methods Eng. **17**, 327–334 (1981)
25. Quey, R., Dawson, P., Barbe, F.: Large-scale 3d random polycrystals for the finite element method: generation, meshing and remeshing. Comput. Methods Appl. Mech. Eng. **200**, 1729–1745 (2011)
26. Reusch, F., Svendsen, B., Klingbeil, D.: Local and non-local Gurson-based ductile damage and failure modelling at large deformation. Eur. J. Mech. A Solids **22**, 779–792 (2003)

27. Steinmann, P., Stein, E.: On the numerical treatment and analysis of finite deformation ductile single crystal plasticity. Comput. Methods Appl. Mech. Eng. **129**, 235–254 (1996)
28. Svendsen, B.: A thermodynamic formulation of finite-deformation elastoplasticity with hardening based on the concept of material isomorphism. Int. J. Plast. **14**, 473–488 (1998)
29. Tvergaard, V., Needleman, A.: Analysis of the cup-cone fracture in a round tensile bar. Acta Metall. **32**, 157–169 (1984)

eXtended Hybridizable Discontinuous Galerkin (X-HDG) for Void and Bimaterial Problems

Ceren Gürkan, Esther Sala-Lardies, Martin Kronbichler, and Sonia Fernández-Méndez

Abstract A strategy for the Hybridizable Discontinuous Galerkin (HDG) solution of problems with voids, inclusions, free surfaces, and material interfaces is proposed. It is based on an eXtended Finite Element (X-FEM) philosophy with a level-set description of interfaces where the computational mesh is not required to fit the interface (i.e. the boundary). This reduces the cost of mesh generation and, in particular, avoids continuous remeshing for evolving interfaces. Differently to previous proposals for the HDG solution with unfitting meshes, the computational mesh covers the domain in our approach, avoiding extrapolations and ensuring the robustness of the method. The local problem in elements not cut by the interface and the global problem are discretized as usual in HDG. A modified local problem is considered for elements cut by the interface. At every cut element, an auxiliary trace variable on the boundary is introduced, which is eliminated afterwards using interface conditions, keeping the original unknowns and the structure of the local problem solver. The solution is enriched with Heaviside functions in case of bimaterial problems; in case of problems with voids, inclusions, or free surfaces no such enrichment is required. Numerical experiments demonstrate how X-HDG keeps the optimal convergence, superconvergence, and accuracy of HDG with no need of adapting the computational mesh to the interface boundary.

1 Introduction

Among state-of-the-art advanced discretization techniques high-order discontinuous Galerkin methods are nowadays very popular in computational fluid dynamics. DG methods inherit the advantages of finite volume methods such as stability through numerical fluxes and local conservation but they allow for the use of high-order

---

C. Gürkan (✉) • E. Sala-Lardies • S. Fernández-Méndez
Universitat Politècnica de Catalunya, Barcelona, Spain
e-mail: ceren.gurkan@upc.edu; esther.sala-lardies@upc.edu; sonia.fernandez@upc.edu

M. Kronbichler
Technische Universität München, München, Germany
e-mail: kronbichler@lnm.mw.tum.de

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_5

approximations with a straightforward implementation of p-adaptivity. Among DG methods the hybridizable discontinuous Galerkin method (HDG) [4, 5], has been successfully applied to all kinds of problems, specifically in computational fluid dynamics (CFD), see, for instance, [6, 19, 20] for its application to the Stokes and Navier–Stokes equations, or [12, 15, 17] for an efficiency comparison to continuous finite elements (CFE) in the context of elliptic problems and wave propagation. HDG inherits all the advantages of high-order DG methods [2, 14, 18, 21], such as local conservation of quantities of interest, intrinsic stabilization thanks to a proper definition of numerical fluxes at element boundaries, suitability for code vectorization and parallel computation, and straight-forward realization of adaptivity. However, two main characteristics allow HDG to outperform other DG methods for problems involving self-adjoint operators, namely hybridization and superconvergence. The hybridization process allows reducing the degrees of freedom in the final linear system to the nodes in the element faces (sides in 2D), similarly to static condensation in the context of high-order continuous elements, see for instance [12]. On other hand, HDG is based on a particular mixed formulation. When approximating both the primal unknown and its derivatives with polynomials of the same degree k, convergence rates of order $k + 1$ in the $\mathcal{L}_2$ norm are observed for both variables. Therefore, a simple element-by-element postprocessing of the derivatives leads to a superconvergent approximation of the primal variables, with convergence of order $k + 2$ in the $\mathcal{L}_2$ norm. The superconvergent solution can be used to define a simple and efficient error estimator, which then can be used to drive an adaptivity process [11, 13]. When it comes to the efficient solution of problems with moving boundaries and interfaces there is still work to be done. In [23], HDG has been applied to the solution of Stokes interface problems, but always considering computational meshes fitting the interface. In[16], HDG has been used for the solution of elliptic material interface problems where the material interface is defined by the computational mesh. There, the mesh needs to be adapted to properly describe the interface geometry, requiring continuous remeshing in the case of evolving interfaces. Even though a methodology for the solution of elliptic problems with meshes not fitting the boundary has been proposed in [3, 8], extrapolating the solution beyond the elements causes some restrictive requirements on the distance from the computational mesh to the boundary which as well limits the practical applicability of the proposal. An alternative strategy for the HDG solution of interface problems, based on an extended finite element (X-FEM) philosophy, is proposed here. The extended finite element method is a clever strategy to treat, for instance, the discontinuities arising at interfaces. The suitability of the X-FEM strategy for the solution of problems with moving boundaries and interfaces in the context of continuous finite elements along with high-order approximation[1, 9] is well-established, see for instance, the overview paper [10] and the references therein. In X-FEM, interfaces are usually represented as the zero-level set of a signed distance function, and the original computational mesh and unknown structure are maintained, with a proper numerical integration at elements and faces cut by the interface [22]. In the case of material interfaces or cracks, the solution is enriched to represent weak or strong discontinuities across interfaces.

Motivated by the demonstrated efficiency of HDG and X-FEM, this works aims to combine these two frameworks for efficiently solving problems with moving boundaries and interfaces described by level sets. The method, labeled eXtended Hybridizable Discontinuous Galerkin (X-HDG), proposes the efficient solution of problems including voids, inclusions, free surfaces, and material interfaces. The solution is enriched with Heaviside functions in the case of material interfaces and, in the case of weak discontinuities, continuity is weakly imposed, emulating the imposition of continuity across element boundaries in standard HDG. X-HDG inherits the advantages of X-FEM methods i.e., the computational mesh is not required to adapt to the interface. This simplifies and reduces the cost of mesh generation while keeping the computational efficiency, stability, accuracy and optimal convergence of HDG. Differently to [3, 8], the computational mesh always covers the domain and, therefore, no extrapolations are required, leading to a more robust method. Hence, X-HDG keeps the superconvergence and stability properties of standard HDG, but in accordance with an X-FEM philosophy.

Section 2 presents the detailed X-HDG formulation and discretization for void problems with Neumann boundary conditions on the interface (i.e. on the boundary not fitted by the computational mesh). The local problem at elements not cut by the interface and the global problem are discretized as usual in HDG. The discretization of the local problem for cut elements is developed in Sect. 2.2. At every cut element, an auxiliary trace variable on the boundary is introduced, which is eliminated afterwards using the boundary conditions on the interface, keeping the original unknowns. The X-HDG formulation is then extended to problems with Dirichlet interfaces. The application to problems involving both Dirichlet and Neumann interfaces is straightforward. Section 3 explains the X-HDG formulation for bimaterial problems mostly referring to the slight differences in the formulation as compared to standard HDG. Here the solution space is enriched with Heaviside functions to introduce the discontinuity within the element for those elements cut by the interface. Elements not cut by the interface are treated with a standard HDG approach. Finally, numerical examples in Sect. 4 demonstrate the applicability of the method and that X-HDG keeps the accuracy, optimal convergence, and superconvergence of HDG. The application of the developed formulation to problems with evolving boundaries avoids the continuous remeshing to fit the interface, with the corresponding saving in computational time. In addition, no loss of accuracy due to the projection of quantities from one mesh to another occurs in this setting. The additional cost of X-HDG in front of HDG is mainly due to the modification of the numerical quadrature, since elemental computations for cut elements are substantially more expensive than those for standard elements. However, it is worth noting that in practical applications the ratio of cut elements to standard elements is small, and the overcost is negligible compared to the cost of remeshing.

2 X-HDG Formulation for Void Problems

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with an interior boundary $\mathcal{I}$ (also referred to as interface) and an exterior boundary $\partial\Omega^{ext} := \partial\Omega \backslash \mathcal{I}$. The following problem is considered,

$$\begin{aligned} -\nabla \cdot (\nu \nabla u) &= f \text{ in } \Omega, \\ \nu \nabla u \cdot \boldsymbol{n} &= g \text{ on } \mathcal{I}, \\ u &= u_D \text{ on } \partial\Omega^{ext}, \end{aligned} \tag{1}$$

where u is the solution, ν is a material coefficient, f is a given source term, u_D are prescribed values on the exterior boundary, and g is a prescribed flux on the interior boundary, i.e. the voids boundary. Neumann boundary conditions are considered on the interior boundary $\mathcal{I}$. The implementation of Dirichlet boundary conditions on $\mathcal{I}$ is straight-forward, see Remark 2. Here, for simplicity, Dirichlet boundary conditions are considered on the exterior boundary $\partial\Omega^{ext}$; other boundary conditions at $\partial\Omega^{ext}$ do not add any difficulty, since they are implemented as in standard HDG.

The domain Ω is assumed to be covered by a finite element mesh with $\mathtt{n_{el}}$ disjoint elements K_i, such that

$$\overline{\Omega} \subset \bigcup_{i=1}^{\mathtt{n_{el}}} \overline{K}_i, \quad K_i \cap K_j = \emptyset \text{ for } i \neq j, \quad \partial\Omega^{ext} \subset \partial \left[\bigcup_{i=1}^{\mathtt{n_{el}}} \overline{K}_i \right].$$

Note that the mesh fits the exterior boundary $\partial\Omega^{ext}$, but some elements may be cut by the interior boundary $\mathcal{I}$, see Fig. 1. The union of all $\mathtt{n_{fc}}$ faces Γ_i (sides for 2D) intersecting the domain Ω is denoted by

$$\Gamma := \bigcup_{i=1}^{\mathtt{n_{el}}} \left[\partial K_i \cap \overline{\Omega} \right] = \bigcup_{f=1}^{\mathtt{n_{fc}}} \left[\Gamma_f \cap \overline{\Omega} \right].$$

The discontinuous setting of (1) consists of some element-by-element equations and some global ones. The local element-by-element problems correspond to the statement of the PDE in (1) with essential boundary conditions in each element K_i,

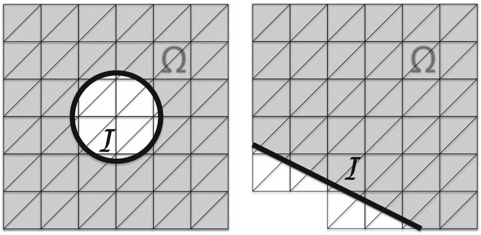


Fig. 1 Two examples of domain with a void: a circular void boundary and a straight interface, $\mathcal{I}$ *black*. The mesh covers the domain Ω (*gray*) and fits the exterior boundary $\partial\Omega^{ext} = \partial\Omega \backslash \mathcal{I}$

that is,

$$\left.\begin{aligned} \nabla\cdot\boldsymbol{q} &= f \text{ in } K_i \\ \boldsymbol{q}+\nu\nabla u &= 0 \text{ in } K_i \\ u &= \hat{u} \text{ on } \partial K_i \end{aligned}\right\} \text{if } \mathcal{I}\cap K_i=\emptyset, \tag{2a}$$

$$\left.\begin{aligned} \nabla\cdot\boldsymbol{q} &= f \text{ in } \Omega_i \\ \boldsymbol{q}+\nu\nabla u &= 0 \text{ in } \Omega_i \\ \boldsymbol{q}\cdot\boldsymbol{n} &= g \text{ on } \mathcal{I}_i \\ u &= \hat{u} \text{ on } \partial\Omega_i\backslash\mathcal{I}_i \end{aligned}\right\} \text{if } \mathcal{I}\cap K_i\neq\emptyset \tag{2b}$$

for $i=1,\dots,\mathtt{n_{el}}$, where, for cut elements,

$$\Omega_i := \Omega\cap K_i,\quad \mathcal{I}_i := \mathcal{I}\cap K_i\,. \tag{3}$$

In (2), two new variables are introduced. The quantity $\boldsymbol{q}$ represents the flux of u in order to formulate the PDE as two first order PDEs, and $\hat{u}$ corresponds to the trace of u at the mesh faces. The local problems have been particularized for elements cut by the interior boundary (2b) and standard elements (2a).

Note that, given the trace $\hat{u}$, which is single valued on the mesh skeleton Γ, the local problems (2) can be solved in each element to determine the solution u and the flux $\boldsymbol{q}$. Thus, the problem now reduces to determine the trace $\hat{u}$, with the so-called *conservativity conditions* (also known as *global equations*), that is, the continuity of the flux across element boundaries

$$[\![\boldsymbol{q}\cdot\boldsymbol{n}]\!]=0 \quad \text{on } \Gamma\backslash\partial\Omega^{ext}, \tag{4}$$

and the boundary condition, equivalent to the exterior boundary condition in (1),

$$\hat{u}=u_D \quad \text{on } \partial\Omega^{ext}. \tag{5}$$

The *jump* $[\![\cdot]\!]$ operator is defined at a face Γ_f as

$$[\![\circledcirc]\!]=\circledcirc_{L(f)}+\circledcirc_{R(f)} \quad \text{on } \Gamma_f,$$

where $R(f)$ and $L(f)$ are numbers of the left and right elements sharing the face, that is, $\Gamma_f=\overline{K_{L(f)}}\cap\overline{K_{R(f)}}$, and the subindex $\circledcirc_i$ denotes the value of function $\circledcirc$ from element K_i. In particular, $[\![\boldsymbol{q}\cdot\boldsymbol{n}]\!]=\boldsymbol{q}_{L(f)}\cdot\boldsymbol{n}_{L(f)}+\boldsymbol{q}_{R(f)}\cdot\boldsymbol{n}_{R(f)}=(\boldsymbol{q}_{L(f)}-\boldsymbol{q}_{R(f)})\cdot\boldsymbol{n}_{L(f)}$.

It is important noting that the continuity of the solution u across Γ is imposed by the Dirichlet boundary condition in the local problems (2) and the fact that $\hat{u}$ is single valued on Γ.

The discretization of the conservativity condition (4) and the local problems (2), with the boundary condition (5), leads to the X-HDG formulation.

The next sections present the details of the X-HDG formulation, stating the discretization of the local problems for standard and cut elements and the discretization of the conservativity condition (4). The local problem at elements not cut by the interface (2a) and the global problem are discretized as usual in HDG [4, 5], as recalled in Sects. 2.1 and 2.3. The discretization of the local problem for cut elements (2b) is developed in Sect. 2.2.

To simplify the presentation, the same notation is used for the numerical approximation, belonging to the usual finite dimensional spaces of piecewise polynomials on elements, and the exact solution, that is u, $\boldsymbol{q}$ and $\hat{u}$.

2.1 Local Problem for Standard Elements

This section recalls the standard HDG local problem on an element K_i not cut by the interface. It corresponds to the discretization of (2a), that is: Given $\hat{u} \in \Lambda^h$, find $u \in \mathcal{P}_k(K_i)$, $\boldsymbol{q} \in [\mathcal{P}_k(K_i)]^d$ where $\mathcal{P}_k$ denotes the space of polynomials of degree less or equal to k, such that

$$\begin{aligned}\int_{K_i} v \boldsymbol{\nabla} \cdot \boldsymbol{q} \, dV + \int_{\partial K_i} \tau v (u - \hat{u}) \, dS = \int_{K_i} v f \, dV \qquad \forall v \in \mathcal{P}_k(K_i),\\ \int_{K_i} \boldsymbol{q} \cdot \boldsymbol{w} \, dV - \int_{K_i} \nu u \boldsymbol{\nabla} \cdot \boldsymbol{w} \, dV + \int_{\partial K_i} \nu \hat{u} \boldsymbol{w} \cdot \boldsymbol{n} \, dS = 0 \ \forall \boldsymbol{w} \in [\mathcal{P}_k(K_i)]^d\end{aligned} \tag{6}$$

The first equation in (6) can be derived from the first equation in (2a) by applying integration by parts, replacing the flux by the numerical flux

$$\hat{\boldsymbol{q}} := \boldsymbol{q} + \tau (u - \hat{u}) \boldsymbol{n}, \tag{7}$$

and undoing the integration by parts. The second equation is obtained from the weak form of the second equation in (2a), applying integration by parts and replacing the boundary condition $u = \hat{u}$ on the element boundary.

Remark 1 The parameter τ, appearing in the definition of the numerical flux (7), is a non-negative stabilization parameter usually taken of order $\mathcal{O}(\nu)$. For each element, it may be taken as a positive constant on all faces, or positive on one arbitrary face and zero on the rest (single face). Both options lead to stable and optimally convergent solutions, with superconvergent post-processed solutions. See for instance [4, 11] for details on the influence of this parameter on the solution behavior.

The discretization of the local problem (6) leads to a system of equations of the form

$$\begin{cases} \mathbf{A}_{uu}^{K_i} \mathbf{u}^i + \mathbf{A}_{uq}^{K_i} \mathbf{q}^i + \mathbf{A}_{u\hat{u}}^{K_i} \boldsymbol{\Lambda}^i = \mathbf{f}_u^{K_i}, \\ \mathbf{A}_{qu}^{K_i} \mathbf{u}^i + \mathbf{A}_{qq}^{K_i} \mathbf{q}^i + \mathbf{A}_{q\hat{u}}^{K_i} \boldsymbol{\Lambda}^i = \mathbf{0}, \end{cases} \tag{8}$$

where $\mathbf{u}^i$ and $\mathbf{q}^i$ are the vectors of nodal values of u and $\boldsymbol{q}$ in element K_i, and $\boldsymbol{\Lambda}^i$ is the vector of nodal values of $\hat{u}$ on the n faces of the element ($n = 3$ for triangles and $n = 4$ for tetrahedra). That is,

$$\boldsymbol{\Lambda}^i := \begin{bmatrix} \hat{\mathbf{u}}^{\mathbf{F}_{i1}} \\ \vdots \\ \hat{\mathbf{u}}^{\mathbf{F}_{in}} \end{bmatrix}, \tag{9}$$

where $\hat{\mathbf{u}}^f$ denotes the nodal values of $\hat{u}$ on face Γ_f, and $\mathbf{F}_{ij}$ is the number of the j-th face of element K_i, Note that the subindices in the $\mathbf{A}$ matrices refer to the space for the weighting function and the test function, respectively.

System (8) can be solved for $\mathbf{u}^i$ and $\mathbf{q}^i$ in each element, obtaining the so-called *local solver* in the element K_i

$$\mathbf{u}^i = \mathbf{U}^{K_i}\boldsymbol{\Lambda}^i + \mathbf{f}_U^{K_i}, \quad \mathbf{q}^i = \mathbf{Q}^{K_i}\boldsymbol{\Lambda}^i + \mathbf{f}_Q^{K_i}, \tag{10}$$

with

$$\begin{bmatrix} \mathbf{U}^{K_i} \\ \mathbf{Q}^{K_i} \end{bmatrix} = -\mathbb{A}^{-1} \begin{bmatrix} \mathbf{A}_{u\hat{u}}^{K_i} \\ \mathbf{A}_{q\hat{u}}^{K_i} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{f}_U^{K_i} \\ \mathbf{f}_Q^{K_i} \end{bmatrix} = \mathbb{A}^{-1} \begin{bmatrix} \mathbf{f}_u^{K_i} \\ \mathbf{0} \end{bmatrix} \tag{11}$$

and

$$\mathbb{A} = \begin{bmatrix} \mathbf{A}_{uu}^{K_i} & \mathbf{A}_{uq}^{K_i} \\ \mathbf{A}_{qu}^{K_i} & \mathbf{A}_{qq}^{K_i} \end{bmatrix}.$$

That is, for each element, the elemental values of the solution, $\mathbf{u}^i$ and $\mathbf{q}^i$, can be explicitly expressed in terms of the trace on its faces, $\boldsymbol{\Lambda}^i$.

2.2 Local Problem for a Cut Element

The X-HDG local problem at an element K_i cut by the interior boundary corresponds to the discretization of (2b), that is: Given $\hat{u} \in \Lambda^h$, find $u \in \mathcal{P}_k(\Omega_i)$, $\boldsymbol{q} \in [\mathcal{P}_k(\Omega_i)]^d$ such that

$$\begin{aligned} &\int_{\Omega_i} v\nabla\cdot\boldsymbol{q}\,dV + \int_{\partial\Omega_i\setminus\mathcal{I}_i} \tau v(u-\hat{u})\,dS + \int_{\mathcal{I}_i} \tau v(u-\tilde{u}^i)\,dS = \int_{\Omega_i} vf\,dV \\ &\int_{\Omega_i} \boldsymbol{q}\cdot\boldsymbol{w}\,dV - \int_{\Omega_i} \nu u\nabla\cdot\boldsymbol{w}\,dV + \int_{\partial\Omega_i\setminus\mathcal{I}_i} \nu\hat{u}\boldsymbol{w}\cdot\boldsymbol{n}\,dS + \int_{\mathcal{I}_i} \nu\tilde{u}^i\boldsymbol{w}\cdot\boldsymbol{n}\,dS = 0 \end{aligned} \tag{12}$$

for all $v \in \mathcal{P}_k(\Omega_i)$ and $\boldsymbol{w} \in [\mathcal{P}_k(\Omega_i)]^d$, where $\Omega_i = \Omega \cap K_i$, and $\tilde{u}^i$ is a new trace variable approximating the trace of the solution on the interface $\mathcal{I}_i = \mathcal{I} \cap K_i$, see Fig. 2. Compared to the weak form for standard elements (6), the X-HDG weak form for a cut element has two additional terms corresponding to integrals along the interface $\mathcal{I}_i$, involving the new trace variable $\tilde{u}^i$.

The discretization of the local problem (12) leads to the system of equations

$$\begin{aligned} \left[\mathbf{A}_{uu}^{\Omega_i} + \mathbf{A}_{uu}^{\mathcal{I}_i}\right]\mathbf{u}^i + \mathbf{A}_{uq}^{\Omega_i}\mathbf{q}^i + \mathbf{A}_{u\hat{u}}^{\Omega_i}\boldsymbol{\Lambda}^i + \mathbf{A}_{u\tilde{u}^i}^{\mathcal{I}_i}\tilde{\mathbf{u}}^i &= \mathbf{f}_u^{\Omega_i}, \\ \mathbf{A}_{qu}^{\Omega_i}\mathbf{u}^i + \mathbf{A}_{qq}^{\Omega_i}\mathbf{q}^i + \mathbf{A}_{q\hat{u}}^{\Omega_i}\boldsymbol{\Lambda}^i + \mathbf{A}_{q\tilde{u}^i}^{\mathcal{I}_i}\tilde{\mathbf{u}}^i &= \mathbf{0}, \end{aligned} \quad , \tag{13}$$

similar to (8), but with three new matrices corresponding to integrals on the Neumann boundary (marked with the superindex $\mathcal{I}_i$), and the nodal values for the new trace variable, $\tilde{\mathbf{u}}^i$. The local problem is closed by imposing the Neumann boundary condition on $\mathcal{I}_i$, i.e.,

$$\hat{\boldsymbol{q}} \cdot \boldsymbol{n} = g \quad \text{on } \mathcal{I}_i.$$

Using the expression of the numerical flux $\hat{\boldsymbol{q}}$ defined in (7), i.e. $\hat{\boldsymbol{q}} \cdot \boldsymbol{n} = \boldsymbol{q} \cdot \boldsymbol{n} + \tau(u - \tilde{u}^i)$, the weak form of the Neumann condition in $\mathcal{I}^i$ is to find $\tilde{u}^i \in \mathcal{P}_k(\mathcal{I}_i)$ such that

$$\int_{\mathcal{I}_i} \tilde{v} \boldsymbol{q} \cdot \boldsymbol{n} \, dS + \tau \int_{\mathcal{I}_i} \tilde{v} \left(u - \tilde{u}^i\right) dS = \int_{\mathcal{I}_i} \tilde{v} g \, dS \quad \forall \tilde{v} \in \mathcal{P}_k(\mathcal{I}_i). \tag{14}$$

The discretization of the Neumann boundary condition leads to a discrete equation of the form

$$\mathbf{A}_{\tilde{u}^i u}^{\mathcal{I}_i}\mathbf{u}^i + \mathbf{A}_{\tilde{u}^i q}^{\mathcal{I}_i}\mathbf{q}^i + \mathbf{A}_{\tilde{u}^i \tilde{u}^i}^{\mathcal{I}_i}\tilde{\mathbf{u}}^i = \mathbf{g},$$

that allows expressing the new trace values $\tilde{\mathbf{u}}^i$ in terms of elemental values

$$\tilde{\mathbf{u}}^i = \mathbf{T}_u^i \mathbf{u}^i + \mathbf{T}_q^i \mathbf{q}^i - \mathbf{t}^i \tag{15}$$

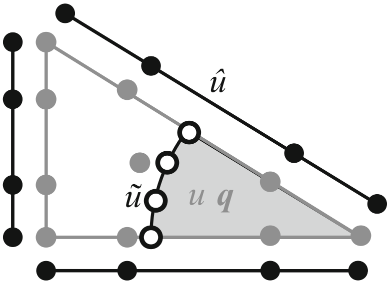


Fig. 2 X-HDG discretization in an element cut by the interior Neumann boundary: *gray* for elemental variables, u and $\boldsymbol{q}$, *black* nodes for trace variable $\hat{u}$ and *white* nodes for the trace on the Neumann boundary $\tilde{u}^i$

with

$$\mathbf{T}^i_u = -[\mathbf{A}^{\mathcal{I}_i}_{\tilde{u}^i\tilde{u}^i}]^{-1}\mathbf{A}^{\mathcal{I}_i}_{\tilde{u}^i u}, \quad \mathbf{T}^i_q = -[\mathbf{A}^{\mathcal{I}_i}_{\tilde{u}^i\tilde{u}^i}]^{-1}\mathbf{A}^{\mathcal{I}_i}_{\tilde{u}^i q}, \quad \mathbf{t}^i = -[\mathbf{A}^{\mathcal{I}_i}_{\tilde{u}^i\tilde{u}^i}]^{-1}\mathbf{g}.$$

Replacing (15) in (13) leads to the final discrete local problem

$$\begin{aligned}\left[\mathbf{A}^{\Omega_i}_{uu} + \mathbf{A}^{\mathcal{I}_i}_{uu} + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{T}^i_u\right]\mathbf{u}^i + \left[\mathbf{A}^{\Omega_i}_{uq} + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{T}^i_q\right]\mathbf{q}^i + \mathbf{A}^{\Omega_i}_{u\hat{u}}\mathbf{\Lambda}^i &= \mathbf{f}^{\Omega_i}_u + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{t}^i\\ \left[\mathbf{A}^{\Omega_i}_{qu} + \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{T}^i_u\right]\mathbf{u}^i + \left[\mathbf{A}^{\Omega_i}_{qq} + \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{T}^i_q\right]\mathbf{q}^i + \mathbf{A}^{\Omega_i}_{q\hat{u}}\mathbf{\Lambda}^i &= \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{t}^i\end{aligned}$$

Now, similarly to (8), this system can be solved for $\mathbf{u}^i$ and $\mathbf{q}^i$, obtaining the *local solver* in the cut element K_i, i.e., Eq. (10) with

$$\begin{bmatrix}\mathbf{U}^{K_i}\\ \mathbf{Q}^{K_i}\end{bmatrix} = -\mathbb{A}^{-1}\begin{bmatrix}\mathbf{A}^{\Omega_i}_{u\hat{u}}\\ \mathbf{A}^{\Omega_i}_{q\hat{u}}\end{bmatrix}, \quad \begin{bmatrix}\mathbf{f}^{K_i}_U\\ \mathbf{f}^{K_i}_Q\end{bmatrix} = \mathbb{A}^{-1}\begin{bmatrix}\mathbf{f}^{\Omega_i}_u + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{t}^i\\ \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{t}^i\end{bmatrix} \tag{16}$$

and

$$\mathbb{A} = \begin{bmatrix}\left[\mathbf{A}^{\Omega_i}_{uu} + \mathbf{A}^{\mathcal{I}_i}_{uu} + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{T}^i_u\right] & \left[\mathbf{A}^{\Omega_i}_{uq} + \mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\mathbf{T}^i_q\right]\\ \left[\mathbf{A}^{\Omega_i}_{qu} + \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{T}^i_u\right] & \left[\mathbf{A}^{\Omega_i}_{qq} + \mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\mathbf{T}^i_q\right]\end{bmatrix}$$

Note that the structure of the local solver is exactly the same as for non-cut elements (10), as the internal trace variable $\tilde{u}^i$ has been isolated and it is not an unknown of the problem anymore.

Remark 2 In case of Dirichlet boundary conditions imposed over the interface $\mathcal{I}$, the X-HDG formulation is straightforward. Once the system in (13) is obtained, the terms $\mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\tilde{\mathbf{u}}^i$ and $\mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\tilde{\mathbf{u}}^i$ are already known so they can simply be moved to the right hand side of the system to obtain a local solver that looks like:

$$\begin{aligned}\left[\mathbf{A}^{\Omega_i}_{uu} + \mathbf{A}^{\mathcal{I}_i}_{uu}\right]\mathbf{u}^i + \mathbf{A}^{\Omega_i}_{uq}\mathbf{q}^i + \mathbf{A}^{\Omega_i}_{u\hat{u}}\mathbf{\Lambda}^i &= \mathbf{f}^{\Omega_i} + \mathbf{f}^{\mathcal{I}_i}_u\\ \mathbf{A}^{\Omega_i}_{qu}\mathbf{u}^i + \mathbf{A}^{\Omega_i}_{qq}\mathbf{q}^i + \mathbf{A}^{\Omega_i}_{q\hat{u}}\mathbf{\Lambda}^i &= \mathbf{f}^{\mathcal{I}_i}_q\end{aligned} \tag{17}$$

where $\mathbf{A}^{\mathcal{I}_i}_{u\tilde{u}^i}\tilde{\mathbf{u}}^i = \mathbf{f}^{\mathcal{I}_i}_u$ and $\mathbf{A}^{\mathcal{I}_i}_{q\tilde{u}^i}\tilde{\mathbf{u}}^i = \mathbf{f}^{\mathcal{I}_i}_q$. Analogous to the Neumann case, (17) can be solved for $\mathbf{u}^i$ and $\mathbf{q}^i$ to obtain the *local solver* for each element and finally the problem can be closed by solving global equation defined in detail in Sect. 2.3. Note that interface problems where some interfaces are subject to Neumann conditions and others are subject to Dirichlet conditions can be naturally covered by selecting the appropriate strategy element-by-element.

2.3 Global Problem

The local problem, both in a standard element or in a cut element, leads to the local solver (10), that computes of the solution in the element, u and $\boldsymbol{q}$, in terms of the trace values at its boundary, $\hat{u}$. Thus, now the problem is reduced to determine the trace nodal values $\{\hat{\mathbf{u}}^f\}_{f=1}^{\mathtt{n_{fc}}}$ on the mesh skeleton Γ. For this purpose the so-called *global problem* is stated, which corresponds to the discretization of the conservativity condition on Γ (4).

Replacing $\boldsymbol{q}$ by the numerical flux (7), the weak form for the trace variable is: Find $\hat{u} \in \Lambda^h$ such that $\hat{u} = u_D$ on $\partial\Omega^{ext}$ and

$$\int_\Gamma \hat{v} [\![\boldsymbol{q} \cdot \boldsymbol{n}]\!] \, dS + 2\tau \int_\Gamma \hat{v} \left(\{u\} - \hat{u} \right) \, dS = 0 \quad \forall \hat{v} \in \Lambda^h,$$

where $\{\cdot\}$ is the mean operator on the faces,

$$\{\circledcirc\} = \frac{1}{2} \left(\circledcirc_{L(f)} + \circledcirc_{R(f)} \right) \quad \text{on } \Gamma_f.$$

As usual in HDG, the discretization of this equation for every face Γ_f leads to an equation of the form

$$\mathbf{A}_{\hat{u}u}^{f,L} \mathbf{u}^{L(f)} + \mathbf{A}_{\hat{u}q}^{f,L} \mathbf{q}^{L(f)} \; + \; \mathbf{A}_{\hat{u}u}^{f,R} \mathbf{u}^{R(f)}(f) + \mathbf{A}_{\hat{u}q}^{f,R} \mathbf{q}^{R(f)} \; + \; \mathbf{A}_{\hat{u}\hat{u}}^{f} \hat{\mathbf{u}}^f = 0. \tag{18}$$

Replacing the local solver (10), for the elements $K_{L(f)}$ and $K_{R(f)}$, in (18) for every face Γ_f, leads to a system of equations involving only the trace variables $\{\hat{\mathbf{u}}^f\}_{f=1}^{\mathtt{n_{fc}}}$.

As usual in an HDG code, the implementation of the method involves a loop over elements. For each element, the matrices and vectors for the local solver (10) are computed, and the contribution to the Eq. (18) is assembled for each face of the element. Once the system is assembled for all elements, and Dirichlet boundary conditions (5) are imposed, the linear system can be solved. Then, given the trace variables $\{\hat{\mathbf{u}}^f\}_{f=1}^{\mathtt{n_{fc}}}$, the interior values $\mathbf{u}^i$ and $\boldsymbol{q}^i$, can be computed for each element using (10). Note that X-HDG keeps the structure of a standard HDG code. The only difference is the modified local problem on cut elements (13), and the corresponding matrices in the local solver (16).

Remark 3 An element-by-element postprocessing procedure allows to compute an X-HDG superconvergent solution. Similarly to standard HDG the superconvergent solution can be computed in every element K_i as the solution of: Find $u^* \in \mathcal{P}_{k+1}(\Omega_i)$ such that

$$\begin{gathered} \int_{\Omega_i} \nu \nabla u^* \cdot \nabla v \, dV = - \int_{\Omega_i} \boldsymbol{q} \cdot \nabla v \, dV \quad \forall v \in \mathcal{P}_{k+1}(\Omega_i), \\ \int_{\Omega_i} u^* \, dV = \int_{\Omega_i} u \, dV, \end{gathered}$$

with $\Omega_i = K_i$ for standard elements, and $\Omega_i = K_i \cap \Omega$ for cut elements. The solution of this element-by-element computation, u^*, converges with order $k + 2$ in the $\mathcal{L}_2$ norm. See [4, 7] for details and other possible computations of a superconvergent solution.

3 X-HDG Formulation for Bimaterial Problems

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain divided into two disjoint subdomains, $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$, $\quad \Omega_1 \cap \Omega_2 = \emptyset$, with an interface $\mathcal{I} = \overline{\Omega_1} \cap \overline{\Omega_2}$. The following bimaterial problem is considered,

$$\begin{aligned} -\nabla \cdot (\nu \nabla u) &= f \text{ in } \Omega_1 \cup \Omega_2, \\ [\![u\boldsymbol{n}]\!] &= 0 \text{ on } \mathcal{I}, \\ [\![\nu \nabla u \cdot \boldsymbol{n}]\!] &= 0 \text{ on } \mathcal{I}, \\ u &= u_D \text{ on } \partial\Omega, \end{aligned} \tag{19}$$

where u is the solution, ν is a material coefficient with discontinuous definition across the interface (that is, $\nu = \nu_i$ in Ω_i for $i = 1, 2$), f is a given source term and u_D are prescribed values at the exterior boundary. The *jump* $[\![\cdot]\!]$ operator is defined at an interface (material interface or, later, faces between elements), using values from the domains to the left and right of the interface, $[\![\circledcirc]\!] = \circledcirc_1 + \circledcirc_2$, always involving the normal vector on the interface.

The partition of Ω in elements K_i induces a new problem equivalent to (19), with some element-by-element equations and some global ones. The local problem at uncut element is the one of standard HDG i.e., equivalent to Eq. (2a), and its discretization can be found in Sect. 2.1. For elements cut by the interface the local problem is

$$\left.\begin{aligned} \nabla \cdot \boldsymbol{q} &= f \text{ in } K_i \backslash \mathcal{I} \\ \boldsymbol{q} + \nu \nabla u &= 0 \text{ in } K_i \backslash \mathcal{I} \\ [\![u\boldsymbol{n}]\!] &= 0 \text{ on } \mathcal{I} \cap K_i \\ [\![\boldsymbol{q} \cdot \boldsymbol{n}]\!] &= 0 \text{ on } \mathcal{I} \cap K_i \\ u &= \hat{u} \text{ on } \partial K_i \end{aligned}\right\} \text{ if } \mathcal{I} \cap K_i \neq \emptyset, \tag{20}$$

and its X-HDG discretization is stated in Sect. 3.1.

The problem is closed with the conservativity condition (4), corresponding to the imposition of continuity across element boundaries, and the boundary condition $\hat{u} = u_D$ on $\partial\Omega$. Its weak form is the same as for standard HDG, see Sect. 2.1.

The following discrete spaces for elemental variables, u and $\boldsymbol{q}$, and for trace variables, $\hat{u}$, are considered,

$$\begin{aligned} \mathcal{V}^h &:= \big\{ v \in \mathcal{L}^2(\Omega) \ : v|_{K_i} \in \mathcal{P}_k(K_i) && \text{if } K_i \cap \mathcal{I} = \emptyset, \\ & \qquad\qquad\qquad\quad v|_{K_i} \in \mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i) && \text{if } K_i \cap \mathcal{I} \neq \emptyset \big\}, \\ \Lambda^h &:= \big\{ \hat{v} \in \mathcal{L}^2(\Gamma) \ : \hat{v}|_{\Gamma_i} \in \mathcal{P}_k(\Gamma_i) && \text{if } \Gamma_i \cap \mathcal{I} = \emptyset, \\ & \qquad\qquad\qquad\quad \hat{v}|_{\Gamma_i} \in \mathcal{P}_k(\Gamma_i) \oplus H\mathcal{P}_k(\Gamma_i) && \text{if } \Gamma_i \cap \mathcal{I} \neq \emptyset \big\}, \end{aligned} \tag{21}$$

where H is a Heaviside function that enriches the approximation in cut elements and on cut faces. The Heaviside function can be defined, for instance, as

$$H = \begin{cases} \ \ 1 & \text{in } \Omega_1 \\ -1 & \text{in } \Omega_2 \end{cases}$$

to introduce discontinuities across the interface $\mathcal{I}$.

In every cut element K_i, given the considered enriched space $\mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)$, u is assumed to be discontinuous across the interface $\mathcal{I}$. The continuity of the solution and the normal flux across the interface is then imposed as usual in HDG by means of an auxiliary trace variable on the interface, which is eliminated afterwards keeping the original unknowns.

At cut faces, a discontinuous approximation for the trace $\hat{u}$ is also considered but, in this case, there is no need to impose weak continuity on it. In this way the resulting formulation is equivalent to a standard HDG formulation on a mesh where some elements are split (for instance into a triangle and a quadrilateral) to fit the interface, with the usual face-by-face discontinuous approximation of the trace.

This definition ensures that the proposed X-HDG formulation keeps the superconvergence and stability properties of standard HDG accordance with an X-FEM philosophy using the original computational mesh with an enriched approximation at cut elements and faces.

3.1 Local Problem at a Bimaterial Element

The X-HDG local problem at an element K_i cut by the interface corresponds to the discretization of (20), that is: Given $\hat{u} \in \Lambda^h$, find $u \in \mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)$, $\boldsymbol{q} \in [\mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)]^d$ such that

$$\begin{gathered} \int_{K_i \setminus \mathcal{I}_i} v \boldsymbol{\nabla} \cdot \boldsymbol{q} \, dV + \int_{\partial K_i} \tau \nu \ v(u - \hat{u}) \, dS + 2 \int_{\mathcal{I}_i} \tau \{ \nu v (u - \tilde{u}^i) \} \, dS = \int_{K_i \setminus \mathcal{I}_i} v f \, dV \\ \int_{K_i \setminus \mathcal{I}_i} \frac{1}{\nu} \boldsymbol{q} \cdot \boldsymbol{w} \, dV - \int_{K_i \setminus \mathcal{I}_i} u \boldsymbol{\nabla} \cdot \boldsymbol{w} \, dV + \int_{\partial K_i} \hat{u} \boldsymbol{w} \cdot \boldsymbol{n} \, dS + \int_{\mathcal{I}_i} \tilde{u}^i [\![\boldsymbol{w} \cdot \boldsymbol{n}]\!] \, dS = 0 \end{gathered} \tag{22}$$

for all $v \in \mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)$ and $\boldsymbol{w} \in [\mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)]^d$, where $\mathcal{I}_i := K_i \cap \mathcal{I}$ is the interface segment in the element and $\tilde{u}^i$ is a new trace variable approximating the trace of the solution on the interface $\mathcal{I}_i$. The mean operator $\{\cdot\}$ is defined taking values from the left and right domains sharing the interface or the mesh face, $\{\circledcirc\} = \frac{1}{2}(\circledcirc_1 + \circledcirc_2)$.

The X-HDG local problem (22) can be derived following the standard HDG rationale for each one of the domains, $K_i \cap \Omega_1$ and $K_i \cap \Omega_2$, and summing the obtained weak forms. Compared to the weak form for standard elements (6), the X-HDG weak form for a cut element has two additional terms corresponding to integrals along the interface $\mathcal{I}_i$ involving the new trace variable $\tilde{u}^i$ and taking into account the discontinuous nature of the approximation space. These new terms weakly impose the condition $u = \tilde{u}^i$ on $\mathcal{I}_i$, ensuring the weak imposition of continuity of u across the interface.

The local problem is now closed by imposing the conservativity condition across the interface $\mathcal{I}_i$, $[\![\hat{\boldsymbol{q}} \cdot \boldsymbol{n}]\!] = 0 \quad \text{on } \mathcal{I}_i$. The weak form of this conservativity condition in $\mathcal{I}^i$ can be written as: given $u \in \mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)$, $\boldsymbol{q} \in [\mathcal{P}_k(K_i) \oplus H\mathcal{P}_k(K_i)]^d$, find $\tilde{u}^i \in \mathcal{P}_k(\mathcal{I}_i)$ such that

$$\int_{\mathcal{I}_i} \tilde{v} [\![\boldsymbol{q} \cdot \boldsymbol{n}]\!] \, dS + 2\tau \int_{\mathcal{I}_i} \tilde{v} \left(\{\nu u\} - \{\nu\} \tilde{u}^i \right) dS = 0 \quad \forall \tilde{v} \in \mathcal{P}_k(\mathcal{I}_i). \tag{23}$$

Similarly to the void cut element with Neumann boundary conditions, Eq. (23) can be used to isolate $\tilde{\mathbf{u}}^i$ in terms of $\mathbf{u}^i$ and $\mathbf{q}^i$. Then, replacing in (22) a local problem with the same structure as standard HDG (10), but modified matrices, is obtained.

4 Numerical Tests

The performance of the novel X-HDG method is tested on three numerical examples. The Laplace equation with known analytical solution is solved over a square domain with a circular void in Sect. 4.1. Boundary condition on the circular boundary is of Neumann type in the first case and of Dirichlet type in the second. The accuracy and the convergence of X-HDG for void problems is tested and compared to HDG with a mesh adapted to the void boundary.

Section 4.2 demonstrates the performance of X-HDG for bimaterial problems where the domain is divided by a straight interface. On each side of the interface different materials are considered and the corresponding error analysis is done. As usual, the performance of X-HDG is compared with the one of HDG by solving exactly the same problem with HDG settings, i.e, a mesh fitting to the interface.

In all numerical tests, the stabilization parameter is $\tau = 1$ on all faces.

4.1 X-HDG with Voids

The first numerical example assesses the performance of X-HDG with Neumann boundary conditions imposed at the void interface. The Laplace Eq. (19), with $\nu = 1$, is solved over a square domain with a centered circular void with radius 0.41, $\Omega = (-1,1)^2 \backslash B((0,0),0.41)$. Neumann boundary conditions are imposed on the void boundary $\mathcal{I} = \partial B((0,0),0.41)$, and Dirichlet boundary conditions are imposed on the exterior boundary $\partial\Omega\backslash\mathcal{I} = \partial\left((-1,1)^2\right)$. Dirichlet and Neumann values and the source term f are set so that the analytical solution is

$$u(x,y) = \exp\left(0.1\sin(5.1x - 6.2y) + 0.3\cos(4.3x + 3.4y)\right).$$

Figure 3 shows the computational domain and the analytical solution.

Figure 4 presents the computational mesh for X-HDG and for standard HDG, respectively. For the X-HDG computation, a regular triangular mesh in the square domain $(-1,1)^2$ is considered, covering the domain Ω and fitting the exterior boundary $\partial\Omega\backslash\mathcal{I}$. A level set function is used to describe the boundary of the void, $\mathcal{I}$. Three kinds of elements appear in the computational mesh. Elements inside the domain (dark gray) are treated as standard HDG elements. For elements cut by the interior boundary $\mathcal{I}$ (light gray), the modified X-HDG local problem is considered, see Sect. 2.2. The elements that are totally inside the void (white) have no contribution to the solution, so they are simply disregarded. The computational mesh for standard HDG is adapted to fit the void boundary, with similar uniform mesh size.

Figure 5 shows the evolution of the $\mathcal{L}_2$ error for decreasing uniform mesh size, for X-HDG and HDG with degrees $k = 2,3,4$, for both the solution and the postprocessed superconvergent solution, see Remark 3. X-HDG keeps optimal convergence with rates close to $k+1$ for the solution, and $k+2$ for the postprocessed solution, with similar levels of accuracy as HDG, while getting rid of adapting the mesh to the void boundary.

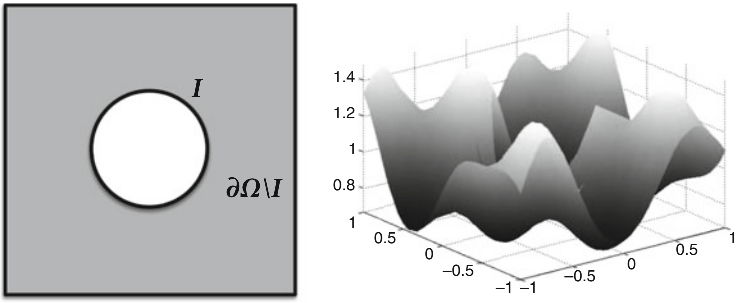


Fig. 3 Circular void example: domain Ω and void boundary $\mathcal{I}$, and analytical solution

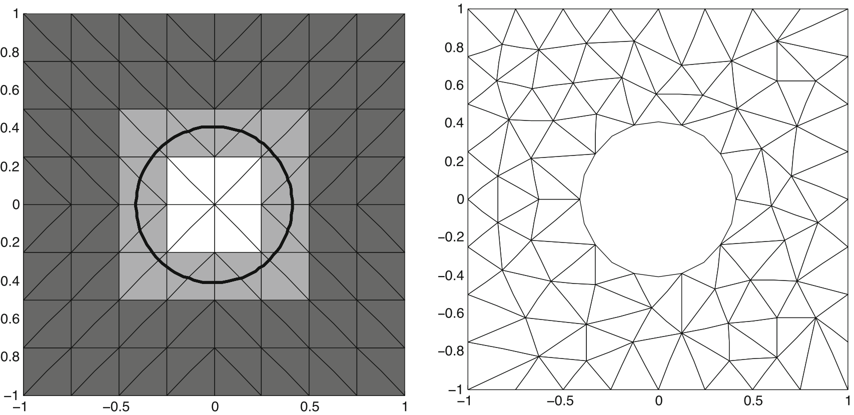


Fig. 4 Circular void example: computational mesh for X-HDG and for standard HDG. The X-HDG mesh is not adapted to the void boundary. Elements in the interior of the domain are colored in *dark gray*. Elements in softer gray are elements cut by the interface $\mathcal{I}$. Elements in white are inside the void, and are not considered in the computation

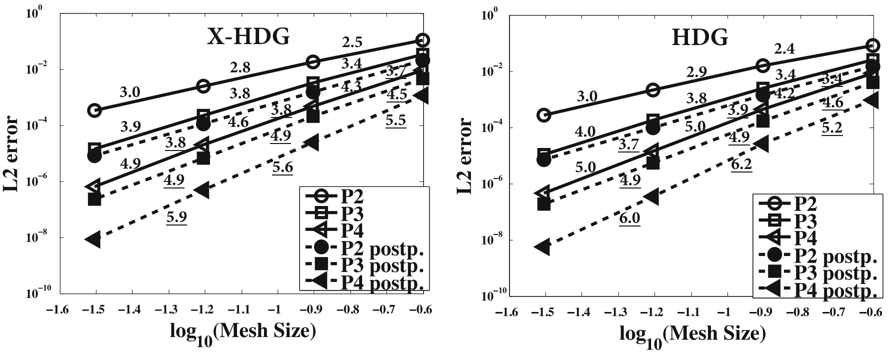


Fig. 5 Neumann circular void example: convergence plots for X-HDG (*left*) and HDG (*right*). The numbers correspond to the slope of each segment, and they are underlined for postprocessed solution

Remark 4 The description of the interface is done using a level set function given by its nodal values on the k-degree mesh used for the solution, and piece-wise $k + 1$ degree level set function is used for the computation of post-processed solution to ensure $k + 2$ convergence rate reached by the postprocessed solution. For standard HDG this is equivalent to requiring a boundary approximation of polynomial degree $k + 1$ properly fitting the description of the boundary for the superconvergent postprocessing.

Now, Dirichlet boundary conditions imposed over the interface $\mathcal{I}$. The domain, the boundary conditions on the exterior boundary, and source term definitions are

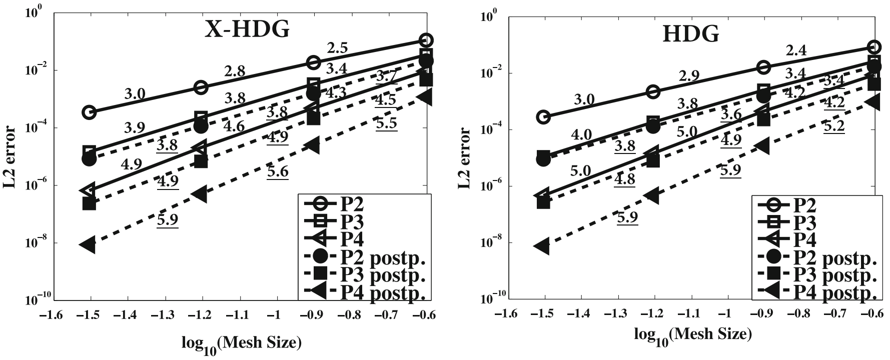


Fig. 6 Dirichlet circular void example: convergence plots for X-HDG (*left*) and HDG (*right*). The numbers correspond to the slope of each segment, and they are underlined for postprocessed solution

kept same as before. Convergence plots are shown in Fig. 6. Again, a one to one resemblance is observed between the X-HDG and HDG results, keeping optimal convergence properties of HDG with an unfitted computational mesh.

4.2 Bimaterial X-HDG

Finally, the Laplace equation (19) over a square domain divided into two disjoint subdomains by a linear interface is solved in order to demonstrate the efficiency of XHDG formulation for bimaterial problems. The size of our domain is $(-1, 1)^2$ and the interface $\mathcal{I}$ is located at $x = 0.2031$ with the viscosity parameter being $\nu_1 = 1$ in the left side of the interface i.e., at Ω_1 and being $\nu_2 = 2.5$ in Ω_2 as shown in Fig. 7.

Homogenous Neumann boundary conditions are set on the interface $\mathcal{I}$ and homogenous Dirichlet boundary conditions set on the outer boundary $\partial\Omega\backslash\mathcal{I} = \partial\left((-1, 1)^2\right)$ where the analytical solution reads

$$u(x) = \begin{cases} 5x^5 & \text{in } \Omega_1, \\ 2x^5 + A & \text{in } \Omega_2, \end{cases}$$

with jump $A = 3(0.2031)^5$. In order to solve this problem, the procedure explained in Sect. 3 is followed for elements cut by the linear interface. The standard HDG formulation is used for elements not cut by the interface.

Figure 8 shows the convergence history of X-HDG and HDG for the bimaterial interface problem. Starting with an initial mesh with four elements per coordinate direction, we have done four mesh refinements—each refinement doubling the element number on the sides—both for X-HDG and HDG with varying approximation

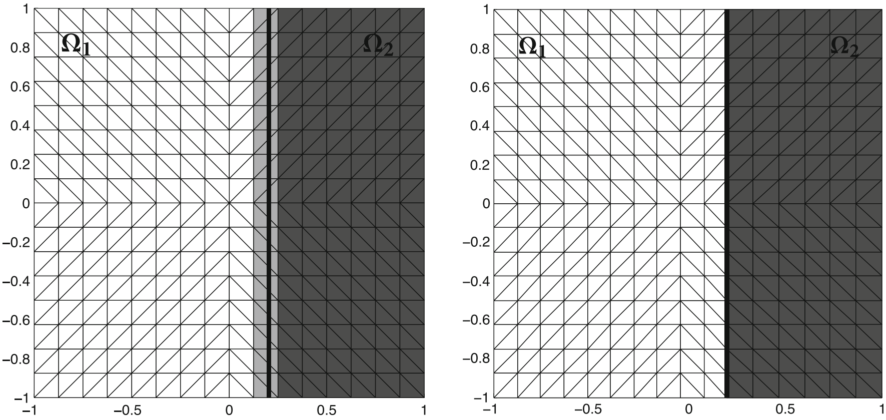


Fig. 7 Bimaterial X-HDG: computational mesh for X-HDG and for standard HDG after two mesh refinements. On the *left*, the X-HDG mesh not adapted to the linear interface boundary $\mathcal{I}$ which shown in *black*. Elements in white are standard elements in domain one whereas elements in *dark gray* are standard elements in domain two. With *light gray*, elements cut by the interface $\mathcal{I}$ are shown. On the *right*, HDG mesh fitting to the linear interface is shown. Elements in domain one are shown in *white* whereas elements in domain two shown in *light gray*

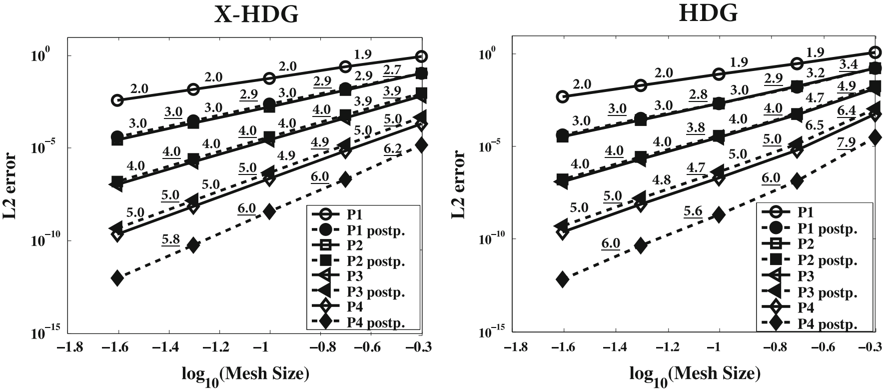


Fig. 8 Bimaterial X-HDG: Convergence history of X-HDG and HDG. Slopes for the solution on each segment are shown in *bold*, slopes for post-processed solution are shown in *bold* and *underlined*

degree between $k = 1$ to $k = 4$. The results verify the X-HDG formulation for bimaterial interface problems when compared to HDG. Over the fine mesh and for higher approximation degree we can see the slight effect of matrix ill-conditioning in both X-HDG and HDG convergence history but this does not prevent any of the methods to reach optimal $k + 1$ and $k + 2$ (super-) convergence rates.

5 Conclusions

A strategy for the HDG solution of boundary value problems with voids, inclusions, free boundaries, and material interfaces has been proposed. It is based on an extended finite element (X-FEM) philosophy with a level-set description of interfaces. With this new strategy, the computational mesh does not need to be adapted to the interface (i.e.; the boundary), simplifying and reducing the cost of mesh generation. In case of evolving interfaces, the cost of continuous re-meshing, and the approximation errors due to the projection from one mesh to another, are avoided.

Differently to previous attempts for the HDG solution with meshes that do not fit the boundary, the computational mesh covers the domain in our method, avoiding extrapolations and ensuring the robustness of the method. The local problem in elements not cut by the interface and the global problem are discretized as usual in HDG. A modified local problem is considered for elements cut by the interface. In every cut element, an auxiliary trace variable on the boundary is introduced, which is eliminated element-by-element using the boundary conditions on the interface, keeping the original unknowns and the structure of the local problem solver. In case of bimaterial problems the solution space is enriched with Heaviside functions to introduce the discontinuity within the element for the elements cut by the interface whereas elements not cut by the interface are treated with standard HDG approach. Although the special treatment for cut elements due to the modified numerical integration causes a slight increase in CPU time, in practical applications the extra cost becomes negligible compared to the cost of mesh adaptation or re-meshing.

The proposed method has been compared to standard HDG on three numerical tests: the solution of a Laplace problem with known analytical solution in a square domain with a circular void with Dirichlet and Neumann boundary conditions, and the solution of the Laplace equation over a square domain with two different materials being divided by a linear interface. Standard HDG computations have been done on a mesh fitting the voids boundaries and material interfaces for comparison with X-HDG that uses a regular mesh covering the domain. The numerical tests assert that X-HDG keeps the HDG optimal convergence rates for the solution, the gradient, and the post-processed super-convergent solution, without the need to adapt the mesh to the boundary. In all tests, similar accuracy is observed for a similar mesh size.

Acknowledgements This work was supported by the DAFOH2 project (Ministerio de Economia y Competitividad, MTM2013-46313-R), and the Erasmus Mundus Joint Doctorate SEED project (European Commission).

References

1. Cheng, K.W., Fries, T.P.: Higher-order XFEM for curved strong and weak discontinuities. Int. J. Numer. Methods Eng. **82**(5), 564–590 (2010)
2. Cockburn, B.: Discontinuous Galerkin methods for computational fluid dynamics. In: Encyclopedia of Computational Mechanics. Fluids, vol. 3, chap. 4. Wiley, New York (2004)
3. Cockburn, B., Solano, M.: Solving convection-diffusion problems on curved domains by extensions from subdomains. J. Sci. Comput. **59**(2), 512–543 (2014)
4. Cockburn, B., Dong, B., Guzmán, J.: A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. Math. Comput. **77**(264), 1887–1916 (2008)
5. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**(2), 1319–1365 (2009)
6. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.J.: Analysis of HDG methods for Stokes flow. Math. Comput. **80**(274), 723–760 (2011)
7. Cockburn, B., Qiu, W., Shi, K.: Conditions for superconvergence of HDG methods for second-order elliptic problems. Math. Comput. **81**(279), 1327–1353 (2012)
8. Cockburn, B., Qiu, W., Solano, M.: A priori error analysis for HDG methods using extensions from subdomains to achieve boundary conformity. Math. Comput. **83**(286), 665–699 (2014)
9. Dreau, K., Chevaugeon, N., Moës, N.: Studied X-FEM enrichment to handle material interfaces with higher order finite element. Comput. Methods Appl. Mech. Eng. **199**(29–32), 1922–1936 (2010)
10. Fries, T.P., Belytschko, T.: The extended/generalized finite element method: An overview of the method and its applications. Int. J. Numer. Methods Eng. **84**(3), 253–304 (2010)
11. Giorgiani, G., Fernández-Méndez, S., Huerta, A.: Hybridizable Discontinuous Galerkin p-adaptivity for wave propagation problems. Int. J. Numer. Methods Fluids **72**(12), 1244–1262 (2013)
12. Giorgiani, G., Modesto, D., Fernández-Méndez, S., Huerta, A.: High-order continuous and discontinuous Galerkin methods for wave problems. Int. J. Numer. Methods Fluids **73**(10), 883–903 (2013)
13. Giorgiani, G., Fernández-Méndez, S., Huerta, A.: Hybridizable Discontinuous Galerkin with degree adaptivity for the incompressible Navier–Stokes equations. Comput. Fluids **98**, 196–208 (2014)
14. Hesthaven, J., Warburton, T.: Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations. J. Comput. Phys. **181**(1), 186–221 (2002)
15. Huerta, A., Angeloski, A., Roca, X., Peraire, J.: Efficiency of high-order elements for continuous and discontinuous Galerkin methods. Int. J. Numer. Methods Eng. **96**(9), 529–560 (2013)
16. Huynh, L., Nguyen, N., Peraire, J., Khoo, B.: A high-order hybridizable discontinuous Galerkin method for elliptic interface problems. Int. J. Numer. Methods Eng. **93**, 183–200 (2012)
17. Kirby, R., Sherwin, S.J., Cockburn, B.: To CG or to HDG: A comparative study. J. Sci. Comput. **51**(1), 183–212 (2011)
18. Montlaur, A., Fernández-Méndez, S., Huerta, A.: Discontinuous Galerkin methods for the Stokes equations using divergence-free approximations. Int. J. Numer. Methods Fluids **57**(9), 1071–1092 (2008)
19. Nguyen, N., Peraire, J., Cockburn, B.: A hybridizable discontinuous Galerkin method for Stokes flow. Comput. Methods Appl. Mech. Eng. **199**(9–12), 582–597 (2010)
20. Nguyen, N.C., Peraire, J., Cockburn, B.: An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations. J. Comput. Phys. **230**(4), 1147–1170 (2011)
21. Peraire, J., Persson, P.O.: The compact discontinuous Galerkin (CDG) method for elliptic problems. SIAM J. Sci. Comput. **30**(4), 1806–1824 (2008)

22. Sala-Lardies, E., Fernández-Méndez, S., Huerta, A.: Optimally convergent high-order X-FEM for problems with voids and inclusions. In: ECCOMAS 2012: 6th European Congress on Computational Methods in Applied Sciences and Engineering. Programme book of abstracts, September 10–14, 2012, Vienna, Austria, pp. 1–14 (2012)
23. Wang, B., Khoo, B.C.: Hybridizable discontinuous Galerkin method (HDG) for Stokes interface flow. J. Comput. Phys. **247**, 262–278 (2013)

Crack Lip Contact Modeling Based on Lagrangian Multipliers with X-FEM

Yuan Jin, Olivier Pierard, Eric Wyart, and Eric Béchet

Abstract The eXtended Finite Element Method (X-FEM), developed intensively in the past 15 years has become a competitive tool for the solution of problems with evolving discontinuities and singularities. In the present study, we focus on the application of X-FEM on frictionless contact problems in the context of fracture mechanics. A promising approach in the literature counting for this problem consists in applying Lagrangian multipliers. Meanwhile, as pointed out in Ji and Dolbow (Int J Numer Methods Eng 61:2508–2535, 2004), a naive choice for Lagrangian multiplier space leads to oscillatory multipliers on the contact surface. This oscillation results from a non-uniform but mesh-dependent inf-sup condition. In this work, we adapt the algorithm proposed in Béchet et al. (Int J Numer Methods Eng 78:931–954, 2009) on crack lip contact by discretizing the displacement field with both scalar and vector tip enrichment functions (Chevaugeon et al., Int J Multiscale Comput Eng 11:597–631, 2013). The influence of the tip enrichment functions on the stability of the formulation is addressed. We show evidences that the vector enrichment functions can improve the conditioning of the problem without jeopardizing the simulation accuracy in the presence of contact.

1 Introduction

The eXtended Finite Element Method (X-FEM) [18], which allows for considering crack geometries within elements, is largely used nowadays to handle crack growth problems. This approach, which uses the Partition of Unity [17], enriches the classical basis of shape functions with the so-called enrichment functions to describe the discontinuity of the displacement field across the crack surface.

In crack propagation simulation, the crack lip contact is important to be considered especially when dealing with complex mechanical loadings, e.g. presence

---

Y. Jin • O. Pierard (✉) • E. Wyart
Cenaero, rue des Frère wright 29, 6041 Gosselies, Belgium
e-mail: yuan.jin@cenaero.be; olivier.pierard@cenaero.be; eric.wyart@cenaero.be

E. Béchet
Université de Liège, Allée de la Découverte 9, B-4000 Liège, Belgium
e-mail: eric.bechet@ulg.ac.be

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_6

of residual stresses generated during an upstreaming manufacturing process and/or compressive loading. A promising route proposed in the literature to solve contact problem in the X-FEM framework is with Lagrangian multipliers. The idea was firstly introduced by Ji and Dolbow[14]. The authors report that a naive discretization of the Lagrangian multipliers violates the inf-sup condition, which determines the stability of the formulation. The problem can be fixed by defining a reduced Lagrangian multiplier space as proposed in [3, 8, 10, 20]. In these works, only Heaviside functions are introduced to represent the jump across the implicit interface. Nevertheless, for crack lip contact, the tip enrichment functions are necessary for the description of the stress singularity around the crack tip and thus for the calculation of the stress intensity factors (SIFs). In other words, these asymptotic functions cannot be omitted in crack propagation simulations. The algorithm of Lagrangian space construction proposed in [3] was applied to model crack lip contact with classical scalar tip enrichment functions on structured meshes in [25]. Meanwhile, the effect of these functions on the stability of the formulation was not discussed by the authors [25]. The present study aims at modeling frictionless crack lip contact on unstructured meshes. Both Heaviside function and tip enrichment functions are used to discretize the displacement field. The influence of the tip enrichment functions on the stability of the formulation is specifically addressed.

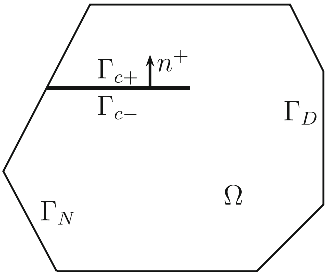
Besides the advantage of relaxing the constraint on finite element mesh, X-FEM is reported to be able to reach the optimal convergence rate with first-order shape functions when the geometrical enrichment strategy is used [2, 15]. However, this enrichment strategy can result in high condition number of the stiffness matrix. The conditioning problem prevents applying this method on complex large industrial problems. Several alternatives are proposed in the literature to address this issue by applying different types of enrichment functions, e.g. the stabilized GFEM [12, 13], the vector enrichment functions [6, 23]. In this work, in addition to classical scalar tip enrichment functions, vector enrichment functions proposed by Chevaugeon et al. [6] are also applied to model crack lip contact. Their performances in terms of simulation accuracy and conditioning are compared.

The paper is organized as follows. We recall at first the contact formulation. Secondly the discretization of the displacement field and the Lagrangian multipliers is presented. The influence of the discretization on the stability of the formulation is then discussed. The fourth section consists in a description of the SIF calculation in the case of crack lip contact. In the last section of the paper, two two-dimensional test cases are applied to analyze the performance of the proposed strategy.

2 Formulation of the Continuous Problem

Hereinafter, bold symbols denote vectors or tensors. We consider a cracked domain Ω in $\mathbb{R}^2$. The two sides of the crack are denoted by Γ_{c+} and Γ_{c-} respectively (Fig. 1). Only normal (frictionless) contact condition is considered between Γ_{c+}

Fig. 1 A cracked domain



and Γ_{c-} in the present study. In the initial configuration, the two sides coincide. $\mathbf{n}$ $(= \mathbf{n}^+ = -\mathbf{n}^-)$ denotes the outward normal unit vector on Γ_{c+}. The body is subjected to surface loads $\mathbf{g} \in \left(L^2(\Gamma_N)\right)^2$. Then under planar small strain assumptions, the problem of homogeneous isotropic linear elasticity consists in finding the displacement field $\mathbf{u} : \Omega \to \mathbb{R}^2$ satisfying

$$\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}) = 0, \qquad \text{in } \Omega, \tag{1}$$

$$\mathbf{u} = 0, \qquad \text{on } \Gamma_D, \tag{2}$$

$$\boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{n} = \mathbf{g}, \qquad \text{on } \Gamma_N, \tag{3}$$

where $\boldsymbol{\sigma} = (\sigma_{ij})$, $1 \leq i, j \leq 2$, stands for the stress tensor field and '$\cdot$'is a contraction over a single index. $\boldsymbol{\sigma}$ is defined by the constitutive law for linear elasticity:

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\varepsilon}, \tag{4}$$

where $\mathbf{C}$ is the Hooke tensor, $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right)/2$ represents the linearized strain tensor field and '$:$'denotes a double contraction.

For a displacement field $\mathbf{u}$, the following notations are adopted to describe the displacement on the contact surface:

$$\mathbf{u}^+ = u_n^+\mathbf{n}^+ + u_t^+\mathbf{t}^+ \quad \text{and} \quad \mathbf{u}^- = u_n^-\mathbf{n}^- + u_t^-\mathbf{t}^-, \tag{5}$$

where $\mathbf{t}$ is the unit tangent vector of Γ_C and $\mathbf{u}^+$ (resp. u^-) is the trace of displacement on Γ_C^+ (resp. on Γ_C^-). Then, the frictionless contact condition reads:

$$\sigma_n(\mathbf{u}) \leq 0, \quad [\![\mathbf{u}]\!] = \left(u_n^+ + u_n^-\right) \leq 0, \quad \sigma_n(\mathbf{u}) \cdot [\![\mathbf{u}]\!] = 0, \tag{6}$$

where $\sigma_n(\mathbf{u})$ is the normal pressure on the contact surface.

In what follows, we adopt classical notations for the functional spaces: $W^{m,p}(\Omega)$ denotes the Sobolev space of functions v for which all components of m^{th} order weak derivative $D^m(v)$ lie in $L^P(\Omega)$. The associated norm is denoted by $\|\cdot\|_{m,p,\Omega}$. Following the convention in the literature [8], we note $H^m(\Omega) := W^{m,2}(\Omega)$, for which the index $p = 2$ is omitted. The components of the solution to problems

(1–6) are then belong to the space $\mathbf{V} := \{\mathbf{v} \in \left(H^1(\Omega)\right)^2, \mathbf{v}|_{\Gamma_D} = 0\}$. As in [1], we introduce the space:

$$W = \{v_n|_{\Gamma_C} : \mathbf{v} \in \mathbf{V}\}. \tag{7}$$

and its dual space W'. We note $H^{1/2}(\Gamma_C)$ the trace space of $H^1(\Omega)$ restricted to Γ_C and $H^{-1/2}(\Gamma_C)$ the dual space of $H^{1/2}(\Gamma_C)$. Then we have $W \subset H^{1/2}(\Gamma_C)$ and $W' \subset H^{-1/2}(\Gamma_C)$. Let $\lambda := \mathbf{n}^T \cdot \boldsymbol{\sigma} \cdot \mathbf{n}$ be the normal surface force on the contact surface. λ then belongs to the space $M = \{\lambda \in W' : \langle \lambda, \psi \rangle_{W',W} \geq 0, \text{ for all } \psi \in W, \psi \leq 0 \text{ on } \Gamma_C\}$. The notation $\langle \cdot, \cdot \rangle_{W',W}$ stands for the duality pairing between W' and W. The weak formulation of the frictionless contact problem then reads, find $\mathbf{u} \in \mathbf{V}$ and $\lambda \in M$ such that

$$\forall \mathbf{u}^* \in \mathbf{V}, \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}\left(\mathbf{u}^*\right) \mathrm{d}\Omega - \int_{\Gamma_C} \lambda [\![\mathbf{u}^*]\!] \mathrm{d}\Gamma = \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{u}^* \mathrm{d}\Gamma. \tag{8}$$

$$\forall \lambda^* \in M, \int_{\Gamma_C} \lambda^* [\![\mathbf{u}]\!] \mathrm{d}\Gamma = 0. \tag{9}$$

3 Discretization

The discretization of the problem (8)–(9) involves a pair of finite element spaces $\mathbf{V}^h \subset \mathbf{V}$ and $M^h \subset M$ leading to the discrete formulation: find $\left(\mathbf{u}^h, \lambda^h\right) \in \mathbf{V}^h \times M^h$ such that

$$\forall \mathbf{u}^{*h} \in \mathbf{V}^h, \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}^h\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}^{*h}\right) \mathrm{d}\Omega - \int_{\Gamma_C} \lambda^h [\![\mathbf{u}^{*h}]\!] \mathrm{d}\Gamma = \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{u}^{*h} \mathrm{d}\Gamma. \tag{10}$$

$$\forall \lambda^{*h} \in M^h, \int_{\Gamma_C} \lambda^{*h} [\![\mathbf{u}^h]\!] \mathrm{d}\Gamma = 0. \tag{11}$$

In this work, we suppose that Ω is meshed with a family of affine triangular meshes $\mathscr{T}_h$, which is assumed to be quasi-uniform and regular. The choice of the $\mathbf{V}^h$ and M^h has to satisfy a uniform inf-sup condition [4] with respect to suitable norms. In the present study, as in [3, 8], the following mesh-dependent L^2 norms are considered:

$$\| \lambda^h \|^2_{-1/2,\Gamma_C} := \sum_{e \in \Gamma_c} h_e \| \lambda^h \|^2_{0,e}, \tag{12}$$

$$\| \mathbf{u}^h \|^2_{1/2,\Gamma_C} := \sum_{e \in \Gamma_c} \frac{1}{h_e} \| \mathbf{u}^h \|^2_{0,e}, \tag{13}$$

where h_e denotes the length of the segment e on Γ_C. Then in the present context, the inf-sup condition reads:

$$\exists c < 0, \text{independent of h}, \forall \lambda_h \in M^h,\ \sup_{\mathbf{u}^h \in \mathbf{V}^h} \frac{\int_{\Gamma_C} \lambda^h [\![\mathbf{u}^h]\!] \mathrm{d}\Gamma}{\| \mathbf{u}^h \|_{1,\Omega}} \geq c h^{1/2} \| \lambda^h \|_{0,\Gamma_C} \tag{14}$$

where $h := \max_e h_e$.

Let U_h and L_h be respectively the vector of the displacement and the Lagrangian multipliers. Then the algebraic form of the problem (10)–(11) reads:

$$\begin{pmatrix} \mathbf{A}_h & \mathbf{B}_h^T \\ \mathbf{B}_h & 0 \end{pmatrix} \begin{pmatrix} U_h \\ L_h \end{pmatrix} = \begin{pmatrix} F_h \\ 0 \end{pmatrix}. \tag{15}$$

To evaluate numerically the inf-sup condition, thanks to the work in [3, 5], the following algebraic equation is introduced:

$$\begin{pmatrix} \mathbf{A}_h & \mathbf{B}_h^T \\ \mathbf{B}_h & \frac{1}{k}\mathbf{M}_h \end{pmatrix} \begin{pmatrix} U_h \\ L_h \end{pmatrix} = \begin{pmatrix} F_h \\ 0 \end{pmatrix} \tag{16}$$

where $\mathbf{M}_h$, a symmetric mass matrix, corresponds to the integral on Γ_C ($\int_{\Gamma_C} \lambda^h \lambda^{*h} \mathrm{d}\Gamma$). Equation (16) is equivalent to Eq. (15) when k tends to infinity. We denote $\beta_{h;min}$ the first non-zero eigenvalue of the following problem:

$$\frac{1}{h} \left(\mathbf{B}_h^T \mathbf{M}_h^{-1} \mathbf{B}_h \right) V_h = \beta_h \mathbf{A}_h V_h. \tag{17}$$

Then the value of c in Eq. (14) is approximated by $\min_{h_i} \beta_{h;min}^{0.5}$.

3.1 Discretization of the Displacement Field

Let $\mathscr{E}_h$ be the set of elements of $\mathscr{T}_h$ completely cut by Γ_C and $\mathscr{V}_h$ the corresponding intersected edges. We denote $\mathscr{N}_h \subset \mathbb{N}$ the nodes of the mesh, $\mathscr{K}_h \subset \mathscr{N}_h$ the nodes of elements in $\mathscr{E}_h$. X-FEM is applied to count for the displacement jump across the crack surface. The discretized displacement field is expressed as follows:

$$\mathbf{u}^h = \sum_{\mathscr{N}_h} N_i(\mathbf{x}) \mathbf{u}_i + \sum_{\mathscr{K}_h} N_k(\mathbf{x}) H(\mathbf{x}) \mathbf{a}_k. \tag{18}$$

The Heaviside-like step function $H(\mathbf{x})$, with a changing sign across the crack, enriches the nodes in $\mathscr{K}_h$. Classically, it is defined thanks to the signed distance function to the crack ϕ, i.e. the so-called normal level set (see [19, 26]):

$$H(\mathbf{x}) = \mathrm{sign}\left(\phi(\mathbf{x})\right). \tag{19}$$

In addition to Heavisde functions, tip enrichment functions are also applied on the set of nodes within a prescribed enrichment radius to the crack tip $\mathscr{J}_h \subset \mathscr{N}_h$ for the purpose of describing the stress singularity. In the present study, two types of tip enrichment functions denoted respectively by scalar enrichment functions and vector enrichment functions are used. For classical scalar enrichment functions, the displacement field is approximated by

$$\mathbf{u}^h = \sum_{\mathscr{N}_h} N_i(\mathbf{x})\mathbf{u}_i + \sum_{\mathscr{K}_h} N_k(\mathbf{x})H(\mathbf{x})\mathbf{a}_k + \sum_{\mathscr{J}_h} N_j(\mathbf{x}) \left(\sum_l f_l(\mathbf{x})\mathbf{b}_{kl} \right), \tag{20}$$

where

$$f = [\sqrt{r}\cos(\theta/2), \sqrt{r}\sin(\theta/2), \sqrt{r}\sin(\theta/2)\sin\theta, \sqrt{r}\cos(\theta/2)\sin(\theta)]. \tag{21}$$

(r, θ) represents the polar coordinate system associated with the crack tip. The vector enrichment functions are introduced in [6]:

$$\mathbf{u}^h = \sum_{\mathscr{N}_h} N_i(\mathbf{x})\mathbf{u}_i + \sum_{\mathscr{K}_h} N_k(\mathbf{x})H(\mathbf{x})\mathbf{a}_k + \sum_{\mathscr{J}_h} \sum_{\alpha} N_j(\mathbf{x})\mathbf{K}_\alpha(\mathbf{x})c_{j\alpha}, \tag{22}$$

where

$$\mathbf{K}_1 = \sqrt{r}\cos(\theta/2)\,(k - \cos(\theta))\,[\mathbf{e}_1(\mathbf{x}) + \mathbf{e}_2(\mathbf{x})]\,, \tag{23}$$

$$\mathbf{K}_2 = \sqrt{r}\sin(\theta/2)\,(k + 2 + \cos(\theta))\,\mathbf{e}_1(\mathbf{x}) + \sqrt{r}\cos(\theta/2)\,(k - 2\cos(\theta))\,\mathbf{e}_2(\mathbf{x}), \tag{24}$$

$$\mathbf{K}_3 = \sqrt{r}\sin(\theta/2)\mathbf{e}_3(\mathbf{x}). \tag{25}$$

As indicated in Eqs. (23)–(25), $\mathbf{K}_\alpha$, $\alpha \in \{1, 2, 3\}$ are defined on the local basis $\mathbf{e}_i$, $i \in \{1, 2, 3\}$. The vector enrichment functions limit the number of degrees of freedom per enriched node to 3 (to be compared to the 12 degrees of freedom per node with scalar enrichment functions). More details of these two types of enrichment functions are provided in [2, 6].

3.2 Discretization of the Lagrangian Multipliers

As in [8, 10], the Lagrangian multiplier components are defined on the parent nodes belonging to $\mathscr{K}_h$. They are based on the same linear shape functions $N_i(\mathbf{x})$ that are used for the displacement field. The approximation space M^h for the multiplier is then:

$$M^h = \{\lambda^h|_\Gamma, \lambda^h \in M'^h\}, \tag{26}$$

where

$$M'^h := \{\mathbf{x} \in \mathscr{E}_h \to \sum_{i \in \mathscr{K}_h} \lambda_i N_i(\mathbf{x}), \lambda_i \in \mathbb{R}\}. \tag{27}$$

With X-FEM, a naive linear $P1 - P1$ interpolation choice for the displacement and Lagrangian multipliers respectively is not stable and results in oscillations [14]. An adequate reduction of the Lagrangian multiplier space can efficiently address the problem [3, 8]. In the algorithm proposed by Béchet et al. [3], a piecewise linear or constant interpolation is adopted for the Lagrangian multipliers. This interpolation, which can be considered as between $P1$ and $P0$, is denoted by $P1^-$ in the present study. Accordingly, combined with the linear interpolation of the displacement, the formulation used is denoted by $P1/P1^-$. Another approach was proposed by Ferté et al. [8]. By releasing some constraint from $P1^-$, the new algorithm of $P1$ multipliers, referred to $P1^*$, verifies $P1^- \subset P1^*$. The formulation is combined with a piecewise quadratic displacement, which is denoted by $P2/P1^*$. Both approaches are proven mathematically to be able to satisfy the inf-sup condition. Roughly speaking, with the constructed Lagrangian multiplier space, it is possible to find $\mathbf{w}^h \in \mathbf{V}^h$ that $\lambda^h = [\![\mathbf{w}^h]\!]$. Then it is sufficient to show $\frac{\int_{\Gamma_C} \lambda^h [\![\mathbf{w}^h]\!] \mathrm{d}\Gamma}{\|\mathbf{w}^h\|_{1,\Omega}} \geq ch^{1/2} \| \lambda^h \|_{0,\Gamma_C}$. We refer to [3, 8] for more details of this demonstration.

In this work, a linear interpolation on the displacement field is adopted. Accordingly, the algorithm of Béchet is applied to construct the Lagrangian multiplier space. It is noteworthy that in [3, 8], the implicit interface completely crosses the simulation domain. However, for contact between crack lips, the contact interface ends within the simulation domain. Therefore, a special treatment is applied on the element containing the crack tip. As illustrated in Fig. 2, the element with nodes $q_i^{ct}, i \in 1, 2, 3$ contains the crack tip. If the edge $q_1^{ct} - q_3^{ct}$ supports an equality relationship, then a complementary equality relationship is imposed so that the degree of freedom of the Lagrangian multiplier defined on q_2^{ct} equals the ones on q_1^{ct} and q_3^{ct}. Otherwise, since either q_3^{ct} or q_1^{ct} is connected to another edge supporting an equality relationship (see [3]), a single equality relationship is imposed among

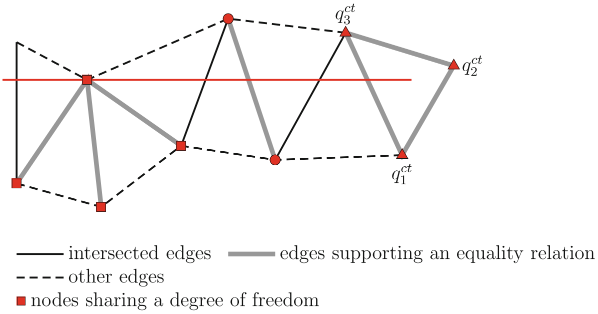


Fig. 2 Restriction algorithm $P1^-$ defined by Béchet et al. [3]

$q_i^{ct}, i \in 1, 2, 3$ so that the node connects the other two nodes. In summary, a $P0$ interpolation is used in this element for Lagrangian multipliers.

Furthermore, in the work of Béchet et al. [3], only the Heavisde function is used for enrichment. In the present study, the tip enrichment functions given in Eqs. (21) and (23)–(25) are also applied to describe the stress singularity around the crack tip. The application of these functions does not violate the inf-sup condition. In fact, by simply imposing the related degrees of freedom ($\mathbf{b}_{kl}$ in Eq. (20) or $c_{j\alpha}$ in Eq. (22)) equal to zero, we can easily find a displacement field $\mathbf{w}^h \in \mathbf{V}^h$ that satisfies $\lambda^h = [\![\mathbf{w}^h]\!]$. Then the demonstration procedure proposed in [3] proves that with this displacement discretization the inf-sup condition is respected.

4 Stress Intensity Factors

An accurate calculation of the stress intensity factors (SIFs) is essential for crack propagation simulation. In the present study, the SIFs are extracted from the X-FEM solution thanks to the interaction integrals in a domain form [7, 16, 24, 28]. Here, we recall the calculation procedure. For this, two states of a cracked body are defined:

- state 1 = present state $\left(u_i^1, \sigma_{ij}^1, \varepsilon_{ij}^1\right)$,
- state 2 = auxiliary state $\left(u_i^2, \sigma_{ij}^2, \varepsilon_{ij}^2\right)$.

Accordingly, the interaction integral $I^{(1,2)}$ can be calculated as follows:

$$I^{(1,2)} = \oint_{\Gamma} \left[W^{(1,2)} \delta_{1j} - \sigma_{ij}^1 \frac{\partial u_i^2}{\partial x_1} - \sigma_{ij}^2 \frac{\partial u_i^1}{\partial x_1} \right] n_j \mathrm{d}\Gamma, \tag{28}$$

where δ is the Kronecker's delta and $W^{(1,2)}$ is the interaction strain energy defined as

$$W^{(1,2)} = \sigma_{ij}^{(1)} \varepsilon_{ij}^{(2)} = \sigma_{ij}^{(2)} \varepsilon_{ij}^{(1)}. \tag{29}$$

The direction x_1 is defined by the local basis associated with the crack tip (see Fig. 3). By defining a weighting function $q(x)$:

$$q = \begin{cases} 1 & \text{on } \Gamma, \\ 0 & \text{on } C_0, \\ \text{arbitrary} & \text{elsewhere,} \end{cases} \tag{30}$$

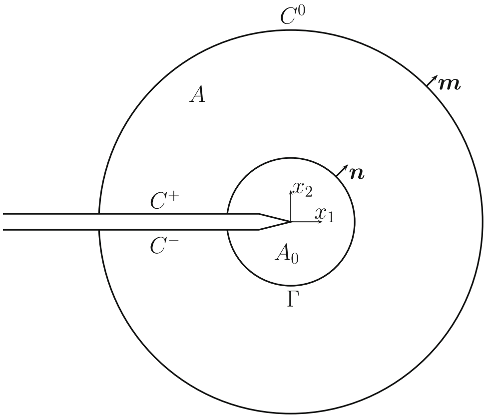


Fig. 3 Contour integral around the crack tip

Equation (28) can be reformulated as:

$$I^{(1,2)} = \int_C \left[-W^{(1,2)}\delta_{1j} + \sigma^1_{ij}\frac{\partial u^2_i}{\partial x_1} + \sigma^2_{ij}\frac{\partial u^1_i}{\partial x_1} \right] q m_j \mathrm{d}C$$
$$- \int_{C_+ + C_-} \left[-W^{(1,2)}\delta_{1j} + \sigma^1_{ij}\frac{\partial u^2_i}{\partial x_1} + \sigma^2_{ij}\frac{\partial u^1_i}{\partial x_1} \right] q m_j \mathrm{d}C, \tag{31}$$

where C is the contour of the Equivalent Domain integral, i.e.

$$C = \Gamma + C_+ + C_- + C_0, \tag{32}$$

and $\mathbf{m}$ is the unit outward normal of C as illustrated in Fig. 3. On Γ, we have $\mathbf{m} = -\mathbf{n}$.

$I^{(1,2)}$ can be written for a general mixed modes (I-II-III) fracture [27]:

$$I^{(1,2)} = \frac{2}{E^*}\left(K^1_I K^2_I + K^1_{II} K^2_{II}\right) + \frac{K^1_{III} K^2_{III}}{\mu}, \tag{33}$$

where $E^* = \frac{E}{1-\nu^2}$ in the present study (plane strain). E and ν denote respectively the Young's modulus and the Poisson ratio. To determine K^1_I, the auxiliary field can be chosen as the asymptotic solution of a pure mode I state so that $K^2_I = 1$, $K^2_{II} = 0$ and $K^2_{III} = 0$. Then K^1_I can be deduced from Eq. (33) as follows:

$$K^1_I = \frac{E^*}{2K^2_I} I^{(1,2)} = \frac{E^*}{2} I^{(1,2)}. \tag{34}$$

For such auxiliary field, $\sigma^{(2)}_{ij} = 0$ for $i = 1, 2$ and $j = 1, 2$ on C_+ and C_-. Accordingly, in case of crack lip contact, the interaction integral $I^{(1,2)}$ can be written as

$$I^{(1,2)} = \int_A \left[-W^{(1,2)}\delta_{1j} + \sigma_{ij}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} + \sigma_{ij}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] \frac{\partial q}{\partial x_j} \mathrm{d}A - \int_{C_+ + C_-} \left[\sigma_{ij}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} \right] q m_j \mathrm{d}C. \tag{35}$$

It is worth emphasizing that different from traction-free cracks, in the case of crack lip contact, the second term in Eq. (35) cannot be ignored. Details of the calculation procedure of the SIFs can be found in [21, 22].

5 Numerical Experiments

In this section, two two-dimensional (2D) frictionless contact problems are used to illustrate the performance of the algorithm described in the previous section. In these two cases, contact forces along the crack are respectively uniform and non-uniform compression. The numerical solutions are compared with analytical values of the Lagrangian multipliers. Plane strain conditions are assumed in both examples. The Young's modulus and the Poisson ratio and are taken respectively as equal to 1 and 0.3.

5.1 Cracked Block Under Uniform Pressure

A uniform pressure (0.01) is imposed on the top surface of a square sample [2 × 2], whereas the bottom right corner is restrained in both directions and the bottom left corner in the vertical (y) one. A crack is inserted into the square sample at $y = 0$ (from $x = -1$ to $x = -0.6$). The test case is illustrated in Fig. 4.

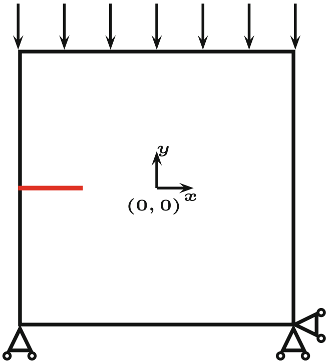


Fig. 4 Uniform pressure imposed on the top surface of a 2D square sample with a crack (*red line*) inside

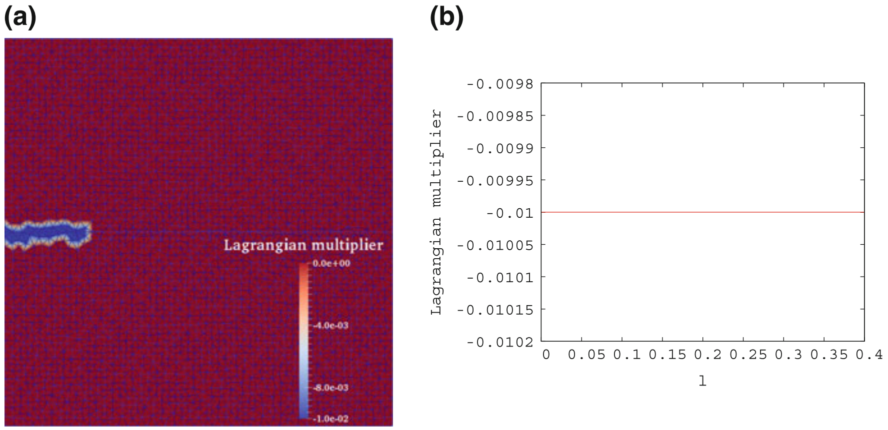


Fig. 5 2D compression test case; displacement field is only enriched with Heaviside functions; (**a**) Lagrangian multiplier in the simulation domain; (**b**) the value of the Lagrangian multiplier on the crack expressed as a function of the distance to the left extremity of the crack

To solve the above problem, firstly as in [3] only Heaviside functions are used as enrichment functions to discretize the displacement field. The problem is solved on an unstructured mesh. 42 equal subdivisions are used along each edge of the square. The mesh size h is thus about 0.05. Applying the previously presented algorithm, the resulted Lagrangian multiplier field is illustrated in Fig. 5. With an adequate Lagrangian multiplier space, the Lagrangian multiplier is constant along the crack and the value corresponds to the pressure imposed on top surface.

Now, for the same problem, the two types of tip enrichment functions are included in the displacement field discretization as indicated by Eqs. (21) and (23)–(25). Geometrical enrichment strategy is applied and the enrichment radius of the tip enrichment functions equals 0.2. However, the application of the full description of the displacement field leads to an oscillation of the Lagrangian multiplier along the crack as illustrated in Fig. 6a. As the same results in terms of Lagrangian multipliers are obtained with the vector enrichment functions, only the results of the scalar functions are shown in this section.

The magnitude of the oscillation is around 1 %. Meanwhile it can be limited to the crack tip by refining the mesh (Fig. 6b). This problem is thus not related to the instability of the formulation. The numerical inf-sup test (Eq. (16)) is performed on this test case. The result is given in Fig. 7. The inf-sup value with a naive Lagrangian multiplier space (with only Heavisde functions) tends to zero, which indicates a violation of the inf-sup condition. On the other hand, with the adequate discretization, regardless the type of tip enrichment functions, the value is more or less constant as the mesh is refined. This observation further shows that the inf-sup condition is respected with tip enrichment functions. The oscillation is very likely to be a numerical error of the finite element resolution.

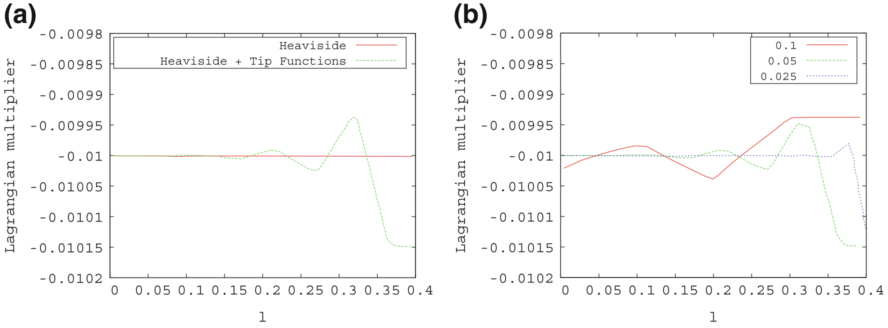


Fig. 6 Oscillation of the Lagrangian multiplier around the crack tip using the scalar tip enrichment functions; (**a**) result on the mesh shown in Fig. 5; (**b**) comparison of the results on different meshes(the mesh size h is indicated in the Figure)

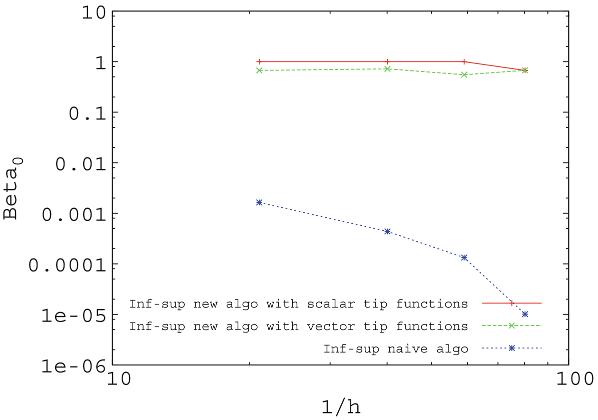


Fig. 7 Numerically computed inf-sup value

Up to now, the integration degree (IntDeg) is set to be 5 on the elements within the enrichment radius. In the present numerical context, the number integration points used for the triangular elements equals $[\mathrm{IntDeg}+3)/2]^2$ while the number of integration points for the linear elements on the crack surface equals $[\mathrm{IntDeg}+1)/2]$ ($[x]$ refers to the integer part of x). The calculation of the stiffness matrix ($\mathbf{A}_h$ in Eq. (15)) contains the integral of the differential of the shape function f_l, which leads to the integral of the singular functions containing a combination of $\{\frac{1}{r}, \frac{1}{\sqrt{r}}, 1, \sqrt{r}, r\}$ multiplied by harmonic functions. Therefore, more integration points are needed to improve the calculation accuracy. As shown in Fig. 8, by doing so, we can considerably decrease the magnitude of the Lagrangian multiplier oscillation around the crack tip.

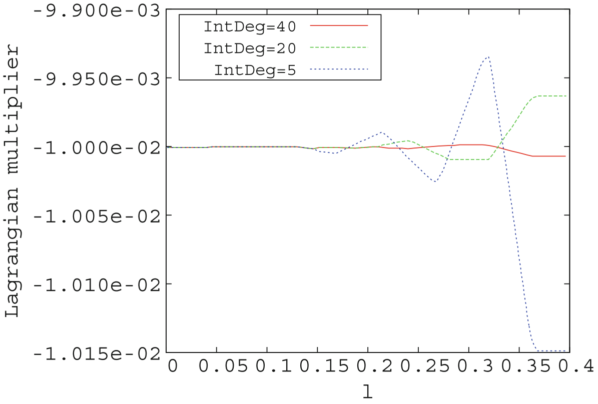


Fig. 8 Lagrangian multiplier value along the crack surface using different integration degrees

Besides the Lagrangian multipliers, the calculation of the SIFs is also analyzed. Since it is a two-dimensional compressive case and only the normal contact is considered, mode I is predominant with respect to mode II. Accordingly, we focus on K_I in this analysis. The discontinuity should be completely removed when the contact condition is fully respected. The theoretical solution of K_I in this case is thus zero. As the local basis is aligned with the global basis and the normal to the crack surface is the axis y, by making use of the asymptotic solution of pure opening mode I as the auxiliary field, Eq. (35) can be further simplified as

$$I^{(1,2)} = \int_A \left[-W^{(1,2)}\delta_{1j} + \sigma_{ij}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} + \sigma_{ij}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] \frac{\partial q}{\partial x_j} \mathrm{d}A + \int_{C_+ + C_-} \left[\lambda \frac{1}{4\mu} \frac{k+1}{\sqrt{2\pi r}} \right] q \mathrm{d}C, \tag{36}$$

where $k = 3 - 4\nu$ under the plane strain assumption. Therefore, the accuracy of the SIF calculation depends on the integral of the singular function $1/\sqrt{r}$. The absolute errors in K_I obtained by applying different integration degrees are expressed as a function of mesh size (h) in Fig. 9. In the present numerical context, for the integral on the crack surface, the crack tip is considered as a node of the surface mesh. Then the relative position of the crack tip with respect to the finite element containing it can have an influence on the integral of the singular function $\frac{1}{\sqrt{r}}$. Here to simplify the discussion, the analysis shown in Fig. 9 is performed exceptionally on structured meshes, which do not fit the crack surface.

As expected, the error in K_I decreases as more integration points and/or smaller elements are used. As indicated in Fig. 9, the convergence rate equals 0.5, which corresponds to the one of the integral of $1/\sqrt{r}$. This observation shows for this problem since the calculation of the Lagrangian multiplier is rather accurate especially when

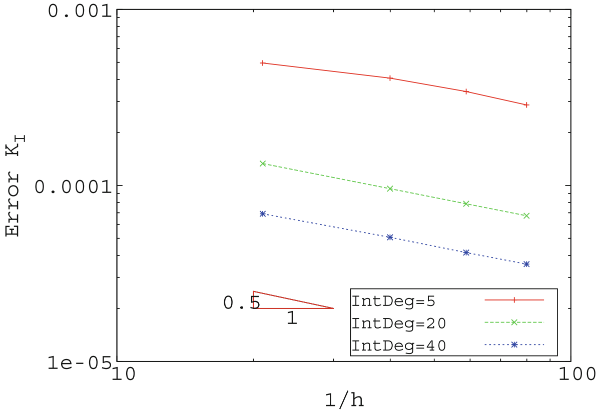


Fig. 9 Evolution of the error in K_I obtained with different integration degrees as a function of mesh size

an elevated degree of integration is used (see Fig. 8), the convergence rate of K_I is mostly determined by the integration accuracy of the singular function $1/\sqrt{r}$ on C_+ and C_-. Addressing this point with a more adapted integration method, e.g. [2, 9] will be a short-term perspective of the present study.

5.2 *Cracked Block Under Non-Uniform Pressure*

In this section, to further test the performance of the proposed algorithm in a more general case, the problem of a crack emanating from a stress raiser is examined: a circular hole in an infinite plate (illustrated in Fig. 10). The exact solution of this problem, given in [11], is expressed in polar coordinates as follows:

$$\sigma_{rr}(r,\theta) = \frac{\sigma_\infty}{2}\left(1-\frac{R^2}{r^2}\right) + \frac{\sigma_\infty}{2}\left(1+\frac{3R^4}{r^4}-\frac{4R^2}{r^2}\right)\cos 2\theta, \tag{37}$$

$$\sigma_{\theta\theta}(r,\theta) = \frac{\sigma_\infty}{2}\left(1+\frac{R^2}{r^2}\right) - \frac{\sigma_\infty}{2}\left(1+\frac{3R^4}{r^4}\right)\cos 2\theta, \tag{38}$$

$$\sigma_{r\theta}(r,\theta) = -\frac{\sigma_\infty}{2}\left(1-\frac{3R^4}{r^4}+\frac{2R^2}{r^2}\right)\sin 2\theta, \tag{39}$$

where R is the hole radius, σ_∞ is the remote compressive load and a is the crack length. The problem is solved on a sequence of uniformly refined unstructured meshes.

To facilitate the implementation of the problem, the exact stress is imposed as Neumann boundary conditions on a semi-circular plate as illustrated by the dash line in Fig. 10. The radius of the hole (R) equals 1, while the radius of the circular

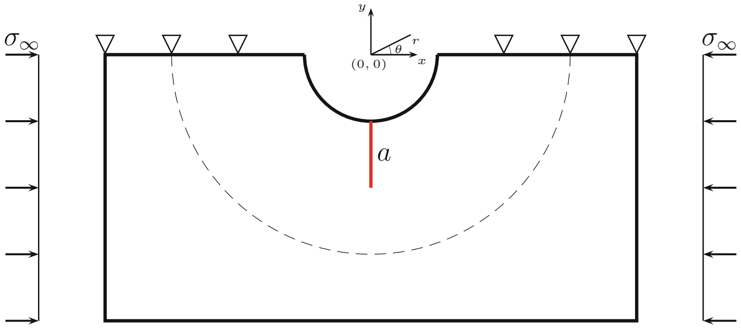


Fig. 10 Uniform pressure imposed on the two sides of an infinite plate with a crack emanating form a circular hole

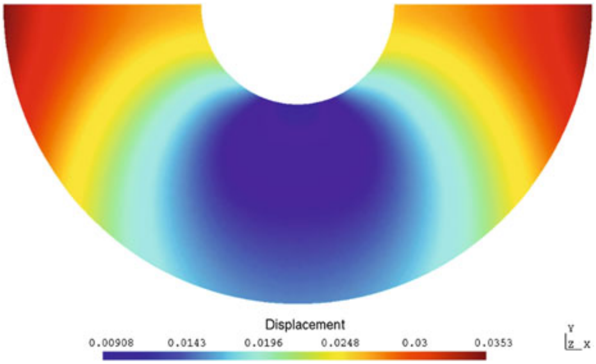


Fig. 11 Solution of the displacement of the infinite plate problem

plate equals 3. The crack length (a) and σ_∞ are respectively set to be 1 and 0.01. Both Heaviside and tip enrichment functions are used to discretize the displacement field. The enrichment radius equals 0.3. To guarantee the accuracy of the Lagrangian multiplier calculation, the degree of integration is set be 20 for the elements within the enrichment radius. The solution of the displacement field is described in Fig. 11.

The evolution of the energy norm error, which is defined by Eq. (40), is described in Fig. 12.

$$\mathrm{Err} = \left(\int_{\Omega} \left(\boldsymbol{\sigma}^{ex} - \boldsymbol{\sigma}^{h} \right)^{T} \mathbf{D} \left(\boldsymbol{\sigma}^{ex} - \boldsymbol{\sigma}^{h} \right) \mathrm{d}\Omega \right)^{0.5} . \tag{40}$$

In addition to the cracked infinite plate problem, a continuous problem without crack in the same configuration has also been solved. With both types of tip enrichment functions, the optimal convergence rate, which is conformed with the continuous problem, is reached.

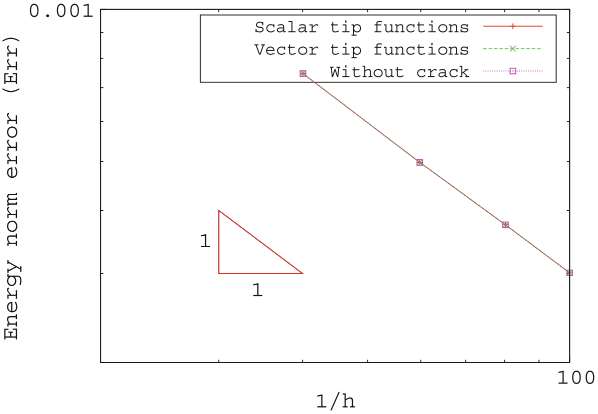


Fig. 12 Evolution of the energy norm error of the infinite plate problem

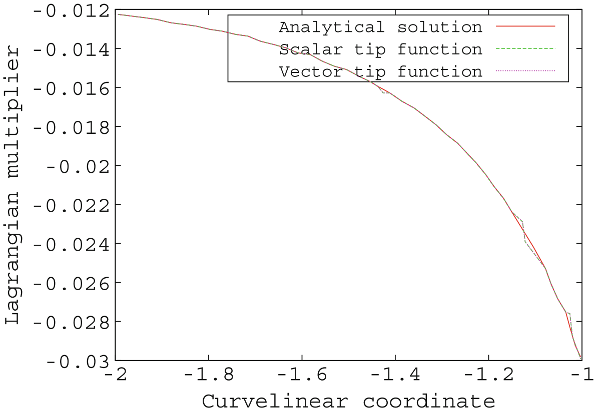


Fig. 13 Profile of the Lagrangian multiplier along the crack obtained on an unstructured mesh ($h = 0.025$) compared with the exact solution

The profile of the Lagrangian multiplier along the crack is compared with the exact solution in Fig. 13. Once again, the numerical solutions with the two types of tip enrichment functions overlap the exact solution. Since the exact stress field is not uniform, to evaluate the calculation of the Lagrangian multiplier, the following L_2 error is introduced:

$$\| e_\lambda \|_{L_2} = \left[\int_{\Gamma_C} (\lambda_{ex} - \lambda_h)^2 \, \mathrm{d}\Gamma \right]^{1/2}. \tag{41}$$

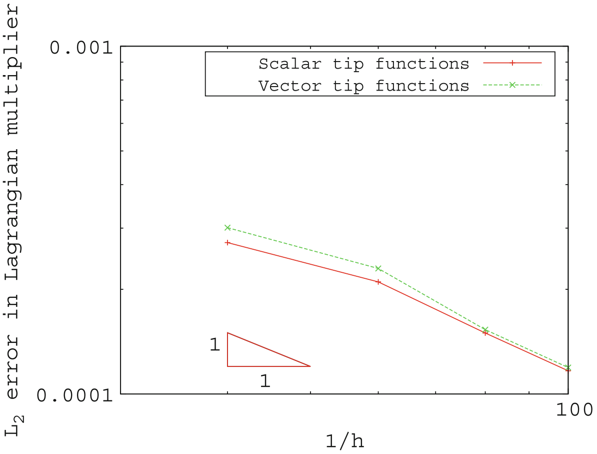


Fig. 14 Evolution of the L_2 error of the Lagrangian multiplier as a function of mesh size (h)

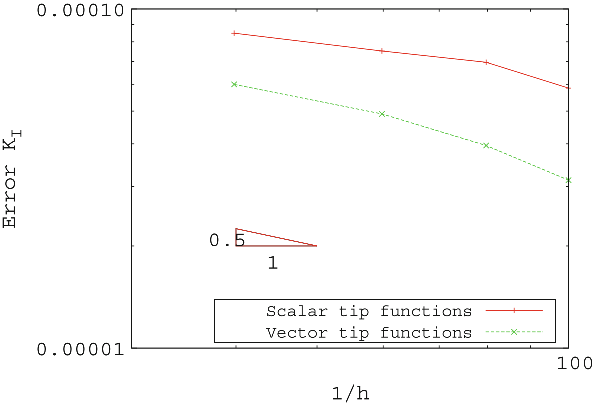


Fig. 15 The error in K_I in the infinite plate problem expressed as a function of mesh size (h)

The evolution of the L_2 error as function of the mesh is described in Fig. 14. As the energy norm error, the errors in the Lagrangian multiplier converge also at the optimal rate, which is equal to 1.

As in the first test case, mode I is predominant with respect to the other mode in the infinite plate problem. The absolute error in K_I is expressed as a function of mesh size in Fig. 15. For this case, K_I is calculated on the unstructured mesh and it does not converge uniformly. Meanwhile, with the two types of enrichment functions, the decrease rate of the absolute error is around 0.5, which shows despite the non-uniform value of the Lagrangian multipliers, the error in K_I is mostly under the influence of the integral accuracy of the singular function $1/\sqrt{r}$ on the crack surface.

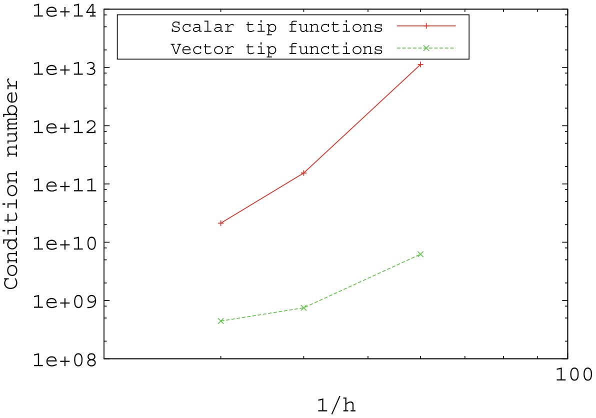


Fig. 16 The evolution of the condition number as a function of mesh size (*h*)

Finally the condition number of the stiffness matrix, defined by the following norm $\| \|$, is expressed as a function of element size (h) in Fig. 16.

$$\| A \| = \max_j \sum_i |A_{ij}|. \tag{42}$$

As shown in [6], The vector enrichment functions improve considerably the conditioning compared to their scalar counterparts. In this case, despite the difference in conditioning, the two types of tip enrichment functions lead to the same overall results. However, this improvement can be essential for more complex problems, e.g. unilateral contact, multi-crack problem, etc.

6 Conclusions

In the present study, the Lagrangian method is applied to consider crack lip contact in an X-FEM context. Following conclusions are drawn from the study:

- The algorithm introduced in [3], which leads to a $P1/P1^-$ formulation, is adapted in the present study by imposing a $P0$ interpolation in the element containing the crack tip to construct the Lagrangian multiplier space along the crack surface.
- In addition to the Heaviside function, tip enrichment functions are included in the displacement discretization. We show theoretically and numerically that the discretization with both scalar and vector tip enrichment functions respects the inf-sup condition.
- The proposed method leads to a precise calculation of Lagrangian multipliers. In the case of constant pressure on the contact surface, slight oscillations (less than 1 %) have been observed around the crack tip. We show evidences that this

oscillation can be considerably reduced by increasing the integration degree. In the case of variable pressure on the contact surface, the vector and scalar enrichment functions result in the same overall results. The application of the vector functions improves considerably the conditioning of the stiffness matrix compared to scalar ones.
- The calculation of SIFs is adapted in the present work to count for the crack lip contact.

The present study was conducted in a two-dimensional context using unstructured meshes, and only frictionless contact is considered. In the future, this work should be at first extended to three-dimensional structures and then to frictional contact. Besides, we observe that the calculation of the Lagrangian multiplier and the stress intensity is tightly related to the integration accuracy. The application of a more adapted integration strategy within the X-FEM framework will be another short-term perspective of this work.

Acknowledgements This work is funded by an SBO Project grant 110070: eSHM with AM of the Agency for Innovation by Science and Technology (IWT). The authors would like to thank C. Friebel and N. Poletz for the insightful discussion about the implementation of the numerical framework.

References

1. Amdouni, S., Moakher, M., Renard, Y.: a stabilized lagrange multiplier method for the enriched finite-element approximation of contact problems of cracked elastic bodies. Comput. Methods Appl. Mech. Eng. **270**, 178–200 (2014)
2. Béchet, E., Minnebo, H., Moës, N., Burgardt, B.: Improved implementation and robustness study of the X-FEM for stress analysis around cracks. Int. J. Numer. Methods Eng. **8**, 1033–1056 (2005)
3. Béchet, E., Moës, N., Wohlmuth, B.: A stable Lagrangian multiplier space for stiff interface conditions within the extended finite element method. Int. J. Numer. Methods Eng. **78**, 931–954 (2009)
4. Brezzi, F., Bathe, K.J.: A discourse on the stability conditions for mixed finite element formulations. Comput. Methods Appl. Mech. Eng. **53**, 27–57 (1990)
5. Chapelle, C., Bathe, K.J.: The inf-sup test. Compos. Struct. **47**, 537–545 (1993)
6. Chevaugeon, N., Moës, N., Minnebo, H.: Improved crack tip enrichment functions and integration for crack modeling using the extended finite element method. Int. J. Multiscale Comput. Eng. **11**, 597–631 (2013)
7. Duflot, M.: A study of the representation of cracks with level sets. Int. J. Numer. Methods Eng. **70**, 1261–1302 (2007)
8. Ferté E., Massin, P., Moës, N.: Interface problems with quadratic X-FEM: design of a stable multiplier space and error analysis. Int. J. Numer. Methods Eng. **11**, 834–870 (2014)
9. Fries, T.P.: A corrected XFEM approximation without problems in blending elements. Int. J. Numer. Methods Eng. **75**, 503–532 (2008)
10. Geniaut, S., Massin, P., Moës, N.: A stable 3D contact formulation using X-FEM. Eur. J. Comput. Mech. **16**, 259–275 (2007)
11. Giner, E., Tur, M., Tarancón, J.E., Fuenmayor, F.J.: Crack face contact in XFEM using a segment to segment approach. Int. J. Numer. Methods Eng. **82**, 1424–1449 (2010)

12. Gupta, V., Duarte, C.A., Babuška, I., Banerjee, U.: A stable and optimally convergent generalized FEM (SGFEM) for linear elastic fracture mechanics. Comput. Methods Appl. Mech. Eng. **266**, 23–39 (2013)
13. Gupta, V., Duarte, C.A., Babuška, I., Banerjee, U.: Stable GFEM (SGFEM): improved conditioning and accuracy of GFEM/XFEM for three-dimensional fracture mechanics. Comput. Methods Appl. Mech. Eng. **289**, 355–386 (2015)
14. Ji, H., Dolbow, J.E.: On strategies for enforcing interfacial constraints and evaluating jump conditions with the extended finite element method. Int. J. Numer. Methods Eng. **61**, 2508–2535 (2004)
15. Laborde, P., Pommier, J., Renard, Y., Salaün, M.: High order extended finite element method for cracked domains. Int. J. Numer. Methods Eng. **64**, 354–381 (2005)
16. Li, F.Z., Shih, C.F., Needleman, A.: A comparison of methods for calculating energy release rates. Eng. Fract. Mech. **21**, 405–421 (1985)
17. Melenk, J.M., Babuška, I.: The partition of unity finite element method: basic theory and applications. Comput. Methods Appl. Mech. Eng. **1–4**, 289–314 (1996)
18. Moës, N., Dolbow, J., Belytschko, T.: A finite element method for crack growth without remeshing. Int. J. Numer. Methods Eng. **1**, 131–150 (1999)
19. Moës, N., Gravouil, A., Belytschko, T.: Non-planar 3D crack growth by the extended finite element and level sets - Part I: mechanical model. Int. J. Numer. Methods Eng. **53**, 2549–2568 (2002)
20. Moës, N., Béchet, E., Tourbier, M.: Imposing Dirichlet boundary conditions in the extended finite element method. Int. J. Numer. Methods Eng. **67**, 1641–1669 (2005)
21. Moran, B., Shih, C.F.: Crack tip and associated domain integrals from momentum and energy balance. Eng. Fract. Mech. **27**, 615–641 (1987)
22. Moran, B., Shih, C.F.: A general treatment of crack tip contour integrals. Int. J. Fract. **35**, 295–310 (1987)
23. Pereira, J.P., Duarte, C.A., Jiao, X.: Three-dimensional crack growth with hp-generalized finite element and face offsetting methods. Comput. Mech. **46**, 431–453 (2010)
24. Shih, C.F., Moran, B., Nakamura, T.: Energy release rate along a three-dimensional crack front in a thermally stressed body. Int. J. Fract. **30**, 79–102 (1986)
25. Siavelis, M., Guiton, M.L.E., Massin, P., Moës, N.: Large sliding contact along branched discontinuities with X-FEM. Comput. Mech. **1**, 201–219 (2012)
26. Stolarska, M., Chopp, D.L., Moës, N., Belyschko T.: Modelling crack growth by level sets in the extended finite element method. Int. J. Numer. Methods Eng. **51**, 943–960 (2001)
27. Wyart, E.: Three-dimensional crack analysis in aeronautical structures using the substructured finite element/extended finite element method. Ph.D Thesis, Universié catholique de Louvain Faculté des Sciences Appliquées (2007)
28. Yau, J.F., Wang, S.S., Corten, H.T.: A mixed-mode crack analysis of isotropic solids using conservation laws of elasticity. J. Appl. Mech. **47**, 335–341 (1980)

Stress Intensity Factors Through Crack Opening Displacements in the XFEM

Markus Schätzer and Thomas-Peter Fries

Abstract The computation of stress intensity factors (SIFs) for two- and three-dimensional cracks based on crack opening displacements (CODs) is presented in linear elastic fracture mechanics. For the evaluation, two different states are involved. An approximated state represents the computed displacements in the solid, which is obtained by an extended finite element method (XFEM) simulation based on a hybrid explicit-implicit crack description. On the other hand, a reference state is defined which represents the expected openings for a pure mode *I*, *II* and *III*. This reference state is aligned with the (curved) crack surface and extracted from the level-set functions, no matter whether the crack is planar or not. Furthermore, as only displacements are fitted, no additional considerations for pressurized crack surfaces are required. The proposed method offers an intuitive, robust and computationally cheap technique for the computation of SIFs where two- and three-dimensional crack configurations are treated in the same manner.

1 Introduction

Simulations in linear elastic fracture mechanics pose a challenging task for numerical methods, as they deal with discontinuities and singularities in solids. For such simulations the efficiency of the standard finite element method (FEM) and boundary element method (BEM) is limited by the $\sqrt{r}$ behaviour of the displacements in the vicinity of the crack tip/front. Special elements at the crack tip/front, like 'quarter-point' elements [13, 20, 21], improve the approach. However, a suitable mesh has to be provided for each crack geometry during propagation which frequently requires a remeshing. The extended finite element method (XFEM) [17, 24] treats these discontinuities and singularities within the elements by additional enrichment functions, wherefore a simple, often structured background mesh is possible with a completely decoupled description of the crack.

M. Schätzer (✉) • T.-P. Fries
Institute of Structural Analysis, Graz University of Technology,
Lessingstr. 25/II, 8010 Graz, Austria
e-mail: schaetzer@tugraz.at; fries@tugraz.at

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_7

In the XFEM, cracks may be defined in different ways. An implicit crack description by means of level-set functions, where the crack geometry is described by the zero level-sets of scalar functions [7, 23]. This description is a standard in the XFEM, however it has the disadvantage that, in general, the update of level-set functions after a propagation step can be cumbersome [7, 12, 22]. A hybrid explicit-implicit crack description [10] combines the advantages of an implicitly defined crack with those of an explicit representation, e.g. by means of straight line segments in two dimensions or flat triangles in three dimensions. An explicit crack description allows a simple update of the crack geometry during the crack propagation as the update of the crack geometry is achieved by adding new segments. This hybrid explicit-implicit description is based on three level-set functions ϕ_j which are derived from the explicitly defined master configuration, see Sect. 2.1. In this work, this second representation of the crack geometry is used.

The behaviour of the fracture including the propagation may be based on stress intensity factors (SIFs). A linear combination of three independent crack modes which are scaled by the SIFs uniquely describes the situation at the crack front. A challenging task is the computation of these scaling factors for non-planar crack geometries with stress-free or loaded crack surfaces. One of the most important technique is based on the interaction integral [6, 18, 26], which is based on the energy release rate G. SIFs are then computed by means of auxiliary fields and the relation of the energy release rate and SIFs. However, for complex three-dimensional crack geometries the evaluation of this integral can be computationally expensive and unstable. In this paper, a more intuitive and computationally cheap method is introduced which allows a straightforward extension to three dimensions also considering mode $I\!I\!I$. This is achieved by observing the displacements, particularly the crack opening displacements (CODs), in the vicinity of the crack tip/front and their comparison with the expected openings for a pure mode I, $I\!I$ and $I\!I\!I$ crack. The fitting of CODs has frequently been done in classical FEM simulations with conforming meshes [5, 20], however, not in an XFEM context.

The paper is organized as follows: Sect. 2 recalls the XFEM approach in linear elastic fracture mechanics and a hybrid explicit-implicit crack description. In Sect. 3, the computation of stress intensity factors (SIFs) by crack opening displacements (CODs) is discussed. Special attention is also given to the location of the fitting approach and the consideration of the crack front. Numerical results in two and three dimensions are presented in Sect. 4. Finally, Sect. 5 concludes this paper with a brief summary.

2 XFEM with a Hybrid Explicit-Implicit Crack Description

In this section, the XFEM approach in linear elastic fracture mechanics with a hybrid explicit-implicit crack description is shortly discussed. Beside of phase-field- or thick level-set-models [4, 15, 19], the XFEM offers a popular approach for linear elastic fracture problems as it allows the approximation of inner-element discontinuities, singularities, kinks and other non-smooth features within elements with optimal accuracy. Therefore, a priori knowledge about the solution characteristics is introduced to the problem using enrichment functions.

Starting point is a domain Ω with the boundary Γ_u where the displacements are prescribed, and the boundary Γ_t where the tractions are prescribed. The domain is cracked by the (curved) crack path/surface Γ_c which may be stress-free or loaded as in applications of hydraulic fracturing. A hybrid explicit-implicit description of the crack is used which is explained in more detail in Sect. 2.1. Cracks feature discontinuous displacements along Γ_c and singular stresses at the crack tip/front. An illustration of the situation is given in Fig. 1 for a two-dimensional case.

Additional enrichment functions which extend the standard finite element approach incorporate the discontinuities and singularities in the displacement and stress fields. For linear elastic fracture problems, the enriched approximation of the XFEM [17] is given by

$$u^h(\mathbf{x}) = \sum_I N_i(\mathbf{x})u_i + \sum_{I^*} N_i(\mathbf{x})\psi_{step}(\mathbf{x})a_i + \sum_{j=1}^{4}\sum_{J^*} N_i(\mathbf{x})\psi^j_{tip}(r,\theta)b^j_i. \tag{1}$$

The first term of the right side represents the standard finite element approach, where N_i are the finite element shape functions. The discontinuities in the displacement field are considered by the additional enrichment function ψ_{step} and the singular stresses at the crack tip/front are considered by the four enrichment functions ψ^j_{tip}. Additional degrees of freedom (a_i, b^j_i) are introduced at the enriched nodes I^* and J^* [17]. In linear elastic fracture mechanics it is standard to use the Heaviside function for the step enrichment ψ_{step} [17] and

$$\psi^j_{tip}(r,\theta) = \left[\sqrt{r}sin\frac{\theta}{2},\ \sqrt{r}cos\frac{\theta}{2},\ \sqrt{r}sin\frac{\theta}{2}sin\theta,\ \sqrt{r}cos\frac{\theta}{2}sin\theta\right] \tag{2}$$

for the crack tip/front enrichment. These enrichment functions are based on a polar coordinate system (r, θ) which is aligned with the tangent at the crack tip/front and has its origin at the crack tip/front. The definition of this coordinate system is based on three level-set functions which is discussed in Sect. 2.1.

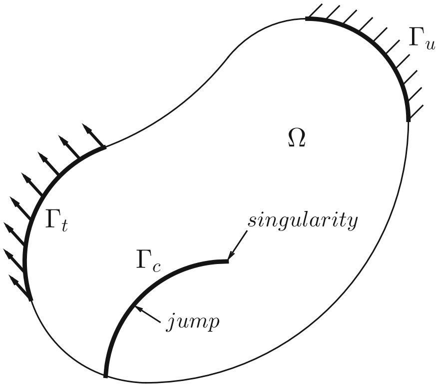


Fig. 1 Situation of a cracked domain in two dimensions

2.1 Hybrid Explicit-Implicit Crack Description

The explicit-implicit crack description is recalled in this section. The aim is to combine an explicit crack description by means of straight line segments in two dimensions or flat triangles in three dimensions, which simplifies the crack update during propagation, and an implicitly defined crack by means of level-set functions. In Fig. 2a, an example for a three-dimensional explicitly defined crack is illustrated. Simulations with the XFEM typically use level-set functions to determine the enrichments. The definition of the enriched nodes (I^* and J^*) is also based on level-set functions, wherefore they are derived from the explicitly defined master configuration. In [10], three level-set functions ϕ_j are introduced which are the basis for the crack location and the enrichment functions in the XFEM. The three level-set functions are defined as follows:

- $\phi_1(\mathbf{x})$ is the (unsigned) distance function to the crack path/surface.
- $\phi_2(\mathbf{x})$ is the (unsigned) distance function to the crack tip/front.
- $\phi_3(\mathbf{x})$ is a signed distance function to the extended crack path/surface.

An illustration of these functions is presented in Fig. 2b–d, where the corresponding level-sets for a three-dimensional crack geometry are illustrated. Furthermore, these functions imply two coordinate systems (r,θ) and (a,b), which are required for the enrichment functions (Sect. 2) and the reference state for the fitting (Sect. 3.2). These two coordinate systems are discussed in the following.

2.1.1 Coordinate System (a, b)

The (a,b)-coordinate system is defined by two scalar functions $a(\mathbf{x})$ and $b(\mathbf{x})$. The correlation of the level-set functions and the coordinates is given by

$$b(\mathbf{x}) = \phi_3(\mathbf{x}) \tag{3}$$

and

$$a^*(\mathbf{x}) = \sqrt{\phi_2(\mathbf{x})^2 - \phi_3(\mathbf{x})^2}. \tag{4}$$

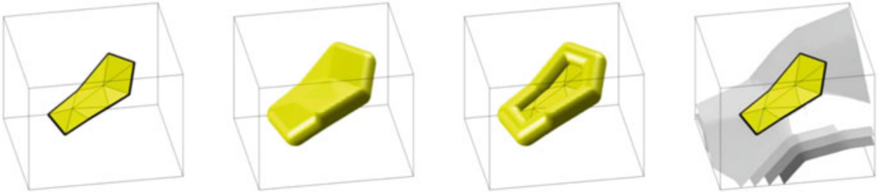


Fig. 2 (**a**) Explicitly defined crack and level-sets: (**b**) $\phi_1 = 1$, (**c**) $\phi_2 = 1$, (**d**) $\phi_3 = 0; 1; 2; 3$ [10]

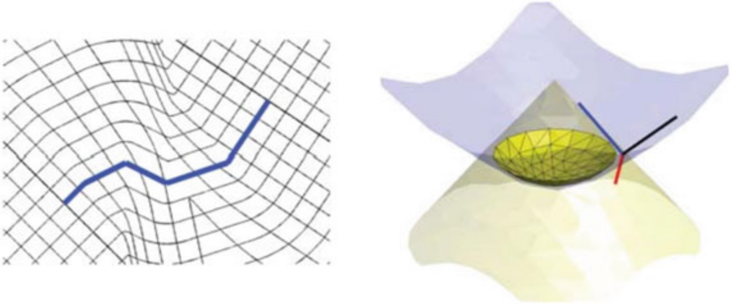


Fig. 3 (**a**) Coordinate system (a, b) in two dimensions and (**b**) zero level-sets of $a(\mathbf{x})$ and $b(\mathbf{x})$ implied by the three level-set functions [10]

As $a^*(\mathbf{x})$ only describes an unsigned distance its sign has to be adapted for $a(\mathbf{x})$, see [10] for further details. Figure 3a shows the (a,b)-coordinate system in two dimensions and (b) the zero level-sets of $a(\mathbf{x})$ and $b(\mathbf{x})$ in three dimensions. This coordinate system simplifies the localisation of the enriched nodes [10]. However, the enrichment functions themselves are based on the polar coordinate system (r, θ), wherefore the relation of these two coordinate systems is discussed in the following section.

2.1.2 Coordinate System (r, θ)

The polar coordinate system (r, θ) can be expressed by the three level-set functions as shown in [10]. A similar setting is used, for example, in [23]. However, for the computation of the expected openings in the reference configuration (Sect. 3.2), it is preferred to express these coordinates by the two scalar functions $a(\mathbf{x})$ and $b(\mathbf{x})$. The radius r is given by

$$r(\mathbf{x}) = \sqrt{a(\mathbf{x})^2 + b(\mathbf{x})^2}. \tag{5}$$

For the definition of θ, several cases must be considered. With the limitation that $b(\mathbf{x}) \neq 0$, the angle $\theta(\mathbf{x})$ is provided by

$$\theta(\mathbf{x}) = \begin{cases} tan^{-1}\left(\frac{b(\mathbf{x})}{a(\mathbf{x})}\right) & \text{for: } a > 0 \\ tan^{-1}\left(\frac{b(\mathbf{x})}{a(\mathbf{x})}\right) + \pi & \text{for: } a < 0 \\ \frac{\pi}{2} & \text{for: } a = 0;\ \ b > 0 \\ \frac{3\pi}{2} & \text{for: } a = 0;\ \ b < 0. \end{cases} \tag{6}$$

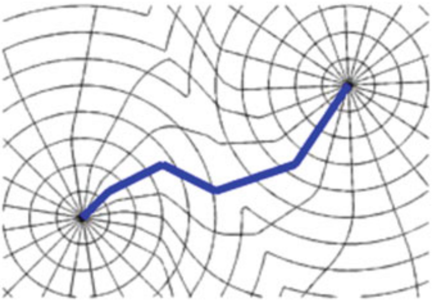


Fig. 4 Coordinate system (r, θ) in 2D implied by the (a, b) coordinate system [10]

A graphical representation of the polar coordinate system (r, θ) in two dimensions is found in Fig. 4.

3 Stress Intensity Factors Through Crack Opening Displacements

The proposed method shows the computation of SIFs by observing the displacement field in the vicinity of the crack tip/front. This method is intuitive and the extension from two to three dimensions is straightforward and treats mode *III* as the other modes. Furthermore, no modification is necessary for curved and loaded cracks. The approximated displacements, which are extracted from the XFEM simulation, are compared with the expected displacements for a pure mode *I*, *II* and *III* crack. That is, two different states are involved which are described in more detail in Sects. 3.1 and 3.2. It is noted that any displacements in the vicinity of the crack tip/front can be used for the computation of SIFs. However, translations by means of rigid body motions have to be extracted from these displacements which complicates the computation. Therefore, we prefer to use relative displacements $\Delta \mathbf{u}$, particularly the crack opening displacements (CODs), as shown in Fig. 5 which consider rigid body motions automatically. The relative displacement $\Delta \mathbf{u}^h_{COD}$ contains the first two modes in two dimensions and all three modes in three dimensions, wherefore all modes can be extracted from this opening. Section 3.3 specifies selected directions where the best results can be expected. In the following sections, the computation of the approximated and expected CODs is discussed in more detail. Additionally, in Sect. 3.3 it is shown how to use these openings for the evaluation of SIFs.

3.1 Approximated State

This section shows the computation of the approximated crack opening $\Delta \mathbf{u}^{h(S)}_{COD}$ of a point S on the crack path/surface. Openings of the point S are computed after solving the linear elastic fracture problem based on Eq. (1) as a post-processing step. The

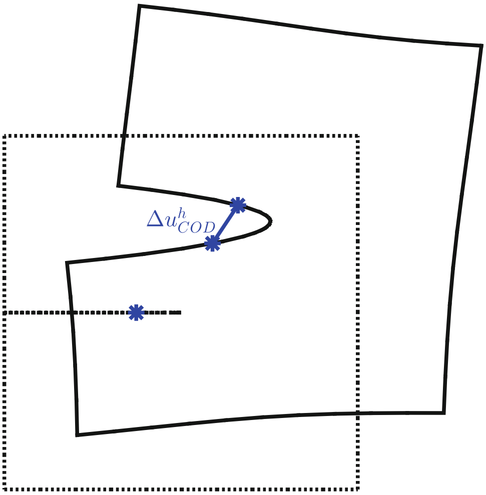


Fig. 5 Crack opening in two dimensions

difficulty here is that the enrichment and shape functions have to be evaluated in a special setting so that CODs are obtained. For that, in a first step a point has to be found on the implicitly described crack path/surface. It is hardly possible to do this in a global setting as the level-set functions are only given at the nodes and are interpolated within the two or three-dimensional elements by the corresponding shape functions [10].

$$\phi_j^h = \sum_i N_i(\mathbf{x})\phi_j(\mathbf{x}_i), \qquad j = 1, 2, 3\,. \tag{7}$$

This interpolation automatically leads to curved zero level-sets in general. However, the following section shows a simplified detection of the zero level-sets. Starting point is a reference quadrilateral or hexahedral element which is called 'master element'. This is decomposed into simplex elements which are called 'sub-elements' and allow for a simple detection of a planar zero level-set. It is noted that a similar setting of the introduced technique is also used for the integration of the domain, see e.g. [9, 11].

3.1.1 Localisation of the Zero Level-Sets

Each potentially cut master element, identified by the condition

$$max[\phi_3(\mathbf{x}_i)] \cdot min[\phi_3(\mathbf{x}_i)] < 0, \tag{8}$$

is subdivided into two linear triangular elements in two dimensions or into six linear tetrahedral elements in three dimensions to ensure that the zero level-set is piecewise straight/planar. Then the level-set functions are interpolated within the sub-elements

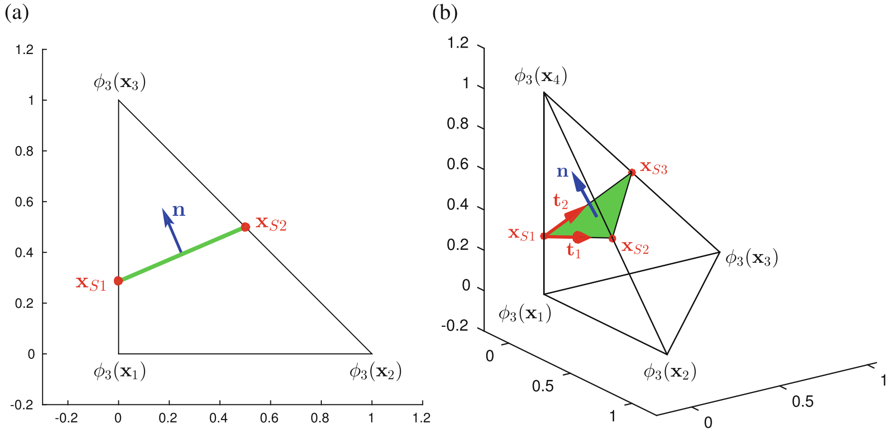


Fig. 6 Zero level-set in reference (**a**) triangular and (**b**) tetrahedral element

by the standard simplex shape functions $N_i^S(\mathbf{x})$. The intersections of the zero level-set with the element edges are marked with $\mathbf{x}_{Si}$, in three dimensions this is: $\mathbf{x}_{Si} = [x_{Si},\ \ y_{Si},\ \ z_{Si}]^T$. Based on the roots on the sub-element edges, the zero level-set is interpolated within the sub-element by the shape functions $N_i^S(\mathbf{x})$. This leads to a straight or flat representation of the zero level-set in the sub-element, as illustrated in Fig. 6. Only three different shapes of the zero level-set are possible: straight lines in two dimensions and flat triangles or quadrilaterals in three dimensions. These straight/flat elements represent the crack path/surface and are, hence, called 'crack elements'. We associate standard finite element shape functions $N_i^C(\mathbf{x})$ to them.

3.1.2 Computation of Approximated CODs

The crack elements are the basis for finding the points $\mathbf{x}_S^j$ on the crack path/surface as all points within these elements can be mapped to the sub-element by

$$\mathbf{x}_S^j = \sum_i N_i^C(\mathbf{x}_j)\mathbf{x}_{Si}. \tag{9}$$

As the displacement field is discontinuous on the crack surface, the point $\mathbf{x}_S^j$ has to be 'splitted' into the two opposite on the crack surface which are infinitesimally close. We call one point S^+ on the positive side and the other one S^- on the negative side. The sign of the sub-domain is defined by the sign of the third level-set function $\phi_3(\mathbf{x})$. This splitting is done in the sub-element by the normalized normal vector $\mathbf{n}_S^j$ as illustrated in Eq. (10).

$$\mathbf{x}_{S\pm}^j = \mathbf{x}_S^j \pm \varepsilon \mathbf{n}_S^j \qquad \text{with} : ||\mathbf{n}_S^j|| = 1 \tag{10}$$

Herein, ε defines the shifting magnitude of the point. It should be small enough so that the splitted points are as close as possible. The normalized normal vector $\mathbf{n}_S^j$ is defined by the cross product of two tangential vectors $\mathbf{t}_1$ and $\mathbf{t}_2$,

$$\mathbf{n}_S^j = \frac{\mathbf{t}_1 \times \mathbf{t}_2}{||\mathbf{t}_1 \times \mathbf{t}_2||}, \tag{11}$$

see Fig. 6. The sign of the normal vector is chosen such that it always points to the positive sub-domain, as shown in Fig. 7 for some example values of the level-set function $\phi_3(\mathbf{x})$ in the corner nodes. There, the master element is illustrated by the dashed line and the sub-element by the solid line. The green line/surface represents the crack element, i.e. the zero level-set within the sub-element. The blue point illustrates a point on the implicitly defined crack path/surface and the blue line its normal vector pointing towards the positive sub-domain.

If the level-set functions and local coordinates (a,b) are known in the two opposite but infinitesimally close points $\mathbf{x}_{S\pm}^j$ their polar coordinates $r(\mathbf{x}_{S\pm}^j)$ and $\theta(\mathbf{x}_{S\pm}^j)$ are determined with Eqs. (5) and (6). Then the approximated displacements $\mathbf{u}_{COD}^{h(S^+)}$ and $\mathbf{u}_{COD}^{h(S^-)}$ are evaluated by Eq. (1). The approximated crack opening $\Delta\mathbf{u}_{COD}^{h(S)}$ is described in the global coordinate system (x, y, z) by the difference of both displacements.

$$\Delta\mathbf{u}_{COD}^{h(S)}(\mathbf{x}_S) = \mathbf{u}_{COD}^{h(S^+)}(\mathbf{x}_{S+}^j) - \mathbf{u}_{COD}^{h(S^-)}(\mathbf{x}_{S-}^j) \tag{12}$$

For the sake of clarity, the index *COD* is not written anymore, instead, the index is used to give information about the current coordinate system, where $(\mathbf{x})$ describes the global coordinate system and $(\mathbf{a})$ the coordinate system mentioned in Sect. 2.1. This information is used for the evaluation of SIFs in Sect. 3.3 to indicate the current coordinate system.

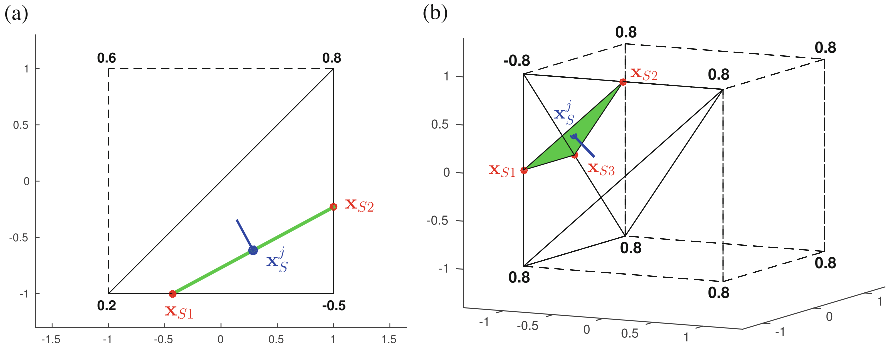


Fig. 7 Master and sub-elements in (**a**) two dimensions and (**b**) three dimensions and the identified straight zero level-set

3.2 Reference State for Pure Mode I, II and III Cracks

In this section, the reference state is discussed which is needed for the evaluation of the expected CODs for a pure mode *I*, *II* and *III* crack, respectively. For straight or planar crack geometries, the definition of the reference configuration can be based on an orthonormal coordinate system which is aligned with the tangent at the crack tip/front. However, in practical applications, cracks are often curved, especially for propagating cracks where mode *II* is dominant. For such problems, the use of an orthonormal coordinate system is not easily justified. We propose the definition of a reference state based on the local coordinate system (a,b) which is also valid for curved cracks. The definition of this coordinate system has been discussed in Sect. 2.1.1. Figure 8a illustrates a two-dimensional curved crack (bold black line) in the (a,b)-coordinate system.

3.2.1 Stress Intensity Approach

A satisfactory description of the crack tip/front behaviour is given by a linear combination of three independent crack modes which are scaled by the SIFs k_I, k_{II} and k_{III}. This approach completely describes the state of the displacements, stresses and strains in the vicinity of the crack tip/front. Under general mixed-mode loadings, the displacements are given in the (a,b,c)-coordinate system by Anderson [1]:

$$\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} \frac{k_I}{2\mu}\sqrt{\frac{r}{2\pi}}cos\frac{\theta}{2}\left(\kappa - 1 + 2sin^2\frac{\theta}{2}\right) + \frac{k_{II}}{2\mu}\sqrt{\frac{r}{2\pi}}sin\frac{\theta}{2}\left(\kappa + 1 + 2cos^2\frac{\theta}{2}\right) \\ \frac{k_I}{2\mu}\sqrt{\frac{r}{2\pi}}sin\frac{\theta}{2}\left(\kappa + 1 - 2cos^2\frac{\theta}{2}\right) - \frac{k_{II}}{2\mu}\sqrt{\frac{r}{2\pi}}cos\frac{\theta}{2}\left(\kappa - 1 - 2sin^2\frac{\theta}{2}\right) \\ \frac{2k_{III}}{\mu}\sqrt{\frac{r}{2\pi}}sin\frac{\theta}{2} \end{bmatrix} \tag{13}$$

$$\text{with} \quad \mu = \frac{E}{2(1+\nu)} \quad \text{and} \quad \kappa = \begin{cases} \frac{3-\nu}{1+\nu} & \text{for plane stress} \\ 3 - 4\nu & \text{for plane strain.} \end{cases}$$

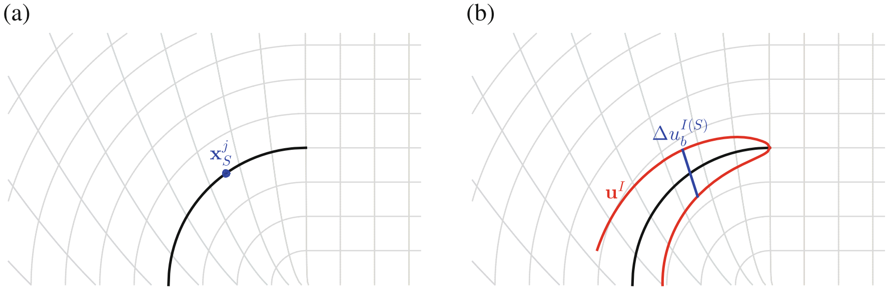


Fig. 8 Curved crack in the (a,b)-coordinate system in a (**a**) closed and (**b**) open setting

Herein, k_I, k_{II} and k_{III} are the SIFs for mode $I - III$, μ the second Lamé-Constant and κ a material parameter. The parameters r and θ describe the polar coordinate system at the crack tip/front, which has been mentioned in Sect. 2.1.2. As Eq. (13) shows, displacements out of the plane only exist if mode III is non zero. However, displacements in direction of a and b consist of a combination of mode I and II for any $0 < |\theta| < \pi$. It is noted that in three dimensions, only plain strain conditions make sense and each point on the crack front has its own stress SIFs.

3.2.2 Computation of Expected CODs

Starting from the definition of the crack in the (a,b)-coordinate system, the expected CODs $\Delta \mathbf{u}_{\mathbf{a}}^{m(S)}$ of the point $\mathbf{x}_S^j$ for a pure mode I, II and III are evaluated. The point $\mathbf{x}_S^j$ and its splitting in two opposite points $\mathbf{x}_{S+}^j$ and $\mathbf{x}_{S-}^j$ is the same as in Sect. 3.1 for which the local coordinates $a(\mathbf{x}_{S\pm}^j)$ and $b(\mathbf{x}_{S\pm}^j)$ are already known. A representation of the undeformed situation is presented in Fig. 8a. By using the relation of the coordinates (a,b) and (r,θ) (Eqs. (5) and (6)) as well as the application of Eq. (13), the expected displacements $\mathbf{u}_{\mathbf{a}}^{m(S^{\pm})}$, with $\mathbf{u}_{\mathbf{a}}^{m(S^{\pm})} = [u_a^{m(S^{\pm})},\ u_b^{m(S^{\pm})},\ u_c^{m(S^{\pm})}]^T$, of the points $\mathbf{x}_{S+}^j$ and $\mathbf{x}_{S-}^j$ are evaluated. Herein, m describes the current crack mode, wherefore, for the expected openings $\mathbf{u}_{\mathbf{a}}^{m(S^{\pm})}$, the stress intensity factor k_m is set to 1 and the others to 0. The expected COD of mode m is then given by the difference of both points

$$\Delta \mathbf{u}_{\mathbf{a}}^{m(S)} = \Delta \mathbf{u}_{\mathbf{a}}^{m(S^{+})}(\mathbf{x}_{S+}^j) - \Delta \mathbf{u}_{\mathbf{a}}^{m(S^{-})}(\mathbf{x}_{S-}^j). \tag{14}$$

In Fig. 8b, the expected opening (red line) for a pure mode I is presented in the (a,b)-coordinate system. Furthermore, the expected COD of the point $\mathbf{x}_S^j$ in direction of b (blue line) is illustrated. Note that the expected openings are given in the (a,b,c)-coordinate system. In the following section, the computation of SIFs based on the evaluated CODs in the approximated and reference state is discussed.

3.3 *Computation of SIFs Using CODs*

A comparison of the approximated and expected openings is only possible if both openings are described in the same coordinate system. Therefore, the approximated CODs, which are given in the global coordinate system (x,y,z), are transformed into the local coordinate system (a,b,c). In two dimensions, this is done based on the

Jacobi-matrix $\mathbf{J}$ of the coordinate transformation, hence,

$$\Delta \mathbf{u}_{\mathbf{a}}^{h} = \mathbf{J} \cdot \Delta \mathbf{u}_{\mathbf{x}}^{h} \qquad \text{with} \quad \mathbf{J} = \begin{bmatrix} a_{,x} & a_{,y} \\ b_{,x} & b_{,y} \end{bmatrix}, \qquad i = 1, 2. \tag{15}$$

It is noted that ∇a and ∇b are only orthonormal for straight/planar cracks. In three dimensions there is no need to explicitly define a third function as ∇c is always normal to ∇a and ∇b. Therefore, ∇c can be computed by the cross-product of ∇a and ∇b.

$$\nabla c = \nabla a \times \nabla b \tag{16}$$

With this information of the third direction, Eq. (15) is straightforwardly extended to the third dimension.

$$\Delta \mathbf{u}_{\mathbf{a}}^{h} = \mathbf{J} \cdot \Delta \mathbf{u}_{\mathbf{x}}^{h} \qquad \text{with} \quad \mathbf{J} = \begin{bmatrix} a_{,x} & a_{,y} & a_{,z} \\ b_{,x} & b_{,y} & b_{,z} \\ c_{,x} & c_{,y} & c_{,z} \end{bmatrix}, \qquad i = 1, 2, 3. \tag{17}$$

Now that both CODs are available in the same coordinate system, the comparison of the approximated and expected CODs leads to the following system of equations:

$$\Delta \mathbf{u}_{\mathbf{a}}^{h} = k_{I} \cdot \Delta \mathbf{u}_{\mathbf{a}}^{I} + k_{II} \cdot \Delta \mathbf{u}_{\mathbf{a}}^{II} + k_{III} \cdot \Delta \mathbf{u}_{\mathbf{a}}^{III}. \tag{18}$$

However, the impact of the individual SIFs to the displacement components of a point which is 'quasi' on the crack surface with: $\theta = \pm(\pi - \varepsilon)$, is quite different as Eq. (13) shows. Herein, k_{I} mainly leads to displacements in direction of b, k_{II} to displacements in direction of a and k_{III} to displacements in direction of c. Therefore, SIFs can be directly computed by

$$k_{I} = \Delta u_{b}^{h} / \Delta u_{b}^{I}; \qquad k_{II} = \Delta u_{a}^{h} / \Delta u_{a}^{II}; \qquad k_{III} = \Delta u_{c}^{h} / \Delta u_{c}^{III}. \tag{19}$$

It remains to specify the location of the considered fitting points and the consideration of the crack front in the context of a numerical simulation.

3.4 Location of the Fitting Points

It is well known that the stress intensity approach is only valid in the vicinity of the crack tip/front, wherefore the maximum distance of the fitting points to the crack tip/front is limited. Numerical results indicate that in two dimensions 10 % of the crack length l_c yield satisfactory results. However, in three dimensions there is no ‘classical’ crack length available, wherefore the maximum distance is limited by 10 % of an *effective* crack length l_c^{eff}. This effective crack length is described by the ratio of the area of the crack surface A_c and the length of the crack front l_{cf}:

$$l_c^{eff} = \frac{A_c}{l_{cf}}. \tag{20}$$

The limitation of the maximum distance of the fitting point $\mathbf{x}_S^j$ to the crack tip/front can then be expressed as follows:

$$r(\mathbf{x}_S^j) \leq \begin{cases} 0.1 \cdot l_c & \text{in 2D} \\ 0.1 \cdot l_c^{eff} & \text{in 3D.} \end{cases} \tag{21}$$

For the evaluation of the partial derivatives of $a(\mathbf{x})$ and $b(\mathbf{x})$ it turned out that it is preferable to use points within sign enriched elements rather than crack-tip enriched elements. In coarse meshes or small crack geometries, it may happen that such points are hard to find. Then, the point $\mathbf{x}_S^j$ is used where the distance $r(\mathbf{x}_S^j)$ is a minimum. Additionally, there exists a wide range of different crack configurations, wherefore it is hardly possible to determine a fixed point on the crack path/surface where the best results are obtained. Therefore, it is proposed to use a number of points in this scope and compute averaged SIFs which also leads to an increased robustness.

In Sect. 3.2.1, it has been mentioned that each point of the crack front has its own SIFs to be determined by Eq. (19). However, for a numerical simulation it is not necessary to know the SIFs of the whole crack front. As [10] shows, it is often sufficient to know the behaviour at the explicitly defined *nodes* on the crack front. Therefore, the following section investigates the situation at the crack front.

3.5 Consideration of the Crack Front

Each point $\mathbf{x}_S^j$ on the implicitly defined crack surface generally describes the behaviour of some related point on the crack front. The localisation of several points which are representative for one special point at the crack front is quite difficult. It is assumed that the point $\mathbf{x}_S^j$ is valid for the point F_j on the crack front where it has a minimum distance, see Fig. 9a. By using an explicit-implicit crack description, a propagation is considered by adding new segments. Therefore, SIFs are needed in the explicitly defined crack front nodes where the direction and distance of the

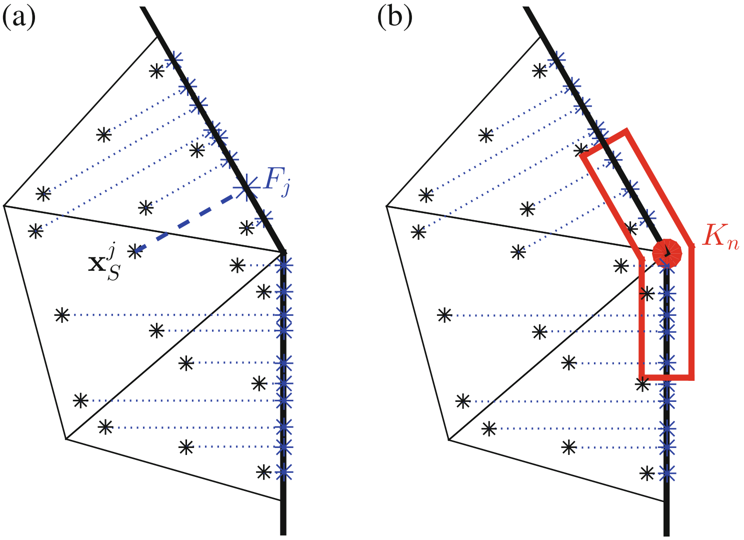


Fig. 9 Three-dimensional crack front: (**a**) Scope of computed SIFs and (**b**) assignment to crack front node

propagation is determined. It can be assumed that the change of the SIFs along the crack front changes slowly for physically justified crack surfaces. Therefore, each evaluated point F_j is assigned to the explicitly defined crack front node K_n where the distance is a minimum, see Fig. 9b. It is noted that the distance of F_j and K_n is also limited by 10 % of the effective crack length, so that only SIFs are considered in the vicinity of the node. However, this limitation is only used for coarsely described crack geometries.

In Fig. 9, a three-dimensional crack front (bold black line) is presented with possible points $\mathbf{x}_S^j$ on the crack surface (black stars). The assignment of these points to the crack front is illustrated by the blue dashed lines and the scope of the crack front node K_n is illustrated by the red line. The SIFs $k_m^{(K_n)}$ of node K_n related to mode m is obtained by the average of all SIFs of the points which are assigned to node K_n. This can be expressed by

$$k_m^{(K_n)} = \sum_{j=1}^{N} \frac{k_m^{(F_j)}}{N}. \tag{22}$$

SIFs are now known for each node of the crack front in the explicit crack description.

4 Numerical Results

In this section, five test cases in linear elastic fracture mechanics are presented in two and three dimensions. The first four examples investigate the behaviour of static planar crack configurations with well known analytical or numerical solutions to show the validity of the proposed method. Herein, SIFs are computed on different meshes and are compared with the expected solutions. In the results, the ratio of the computed and the expected SIFs are plotted over the used number of elements. The first two examples are in two dimensions and refer to externally loaded domains with different boundary conditions and stress-free crack surfaces. Loaded crack surfaces are illustrated in examples three and four, where one is in two and the other in three dimensions. A quasi-static two-dimensional crack propagation is presented in the last example to show the accuracy of the proposed method also for curved cracks. Each mesh is based on bilinear 4-node quadrilateral or trilinear 8-node hexahedral elements. Furthermore, a brittle and isotropic material is used with a Young's modulus $E = 35\,\text{GPa}$ and a Poisson's ratio $\nu = 0.3$. In two dimensions, plane stress conditions are assumed. The interest is only in the computation of SIFs in brittle materials under quasi-static mixed-mode loadings, wherefore no dynamic effects [14] or cohesive models [16] are considered herein.

4.1 Eccentric Three-Point Bending Test

This test case shows an edge cracked three-point bending test with an eccentric load, where $F = 100\,\text{kN}$ and the eccentricity $d = 75\,\text{cm}$. The beam is 600 cm long (l) and 150 cm high (h) and exhibits an $l_c = 75\,\text{cm}$ long initial crack in the middle, see Fig. 10a. The expected SIFs are given [8] by

$$\begin{bmatrix} k_I \\ \\ k_{II} \end{bmatrix} = \begin{bmatrix} \frac{3Fl\sqrt{\pi l_c}}{h^2} \cdot \frac{F_I'}{(1-\frac{l_c}{h})^{\frac{3}{2}}} \\ \\ \frac{3Fl\sqrt{\pi l_c}}{h^2} \cdot F_{II} \end{bmatrix} . \tag{23}$$

With $F_I' = 0.4010$ and $F_{II} = 0.0876$, Eq. (23) leads to the expected SIFs $k_I = 69.64$ and $k_{II} = 5.38$. Starting with a coarse mesh and 390 elements as shown in Fig. 10b, SIFs are computed on seven different meshes, which are generated by a refinement until a fine mesh with 67.470 elements is obtained. A local refinement in the vicinity of the crack allows a limitation of the element number. Figure 10c shows the normalized results of the computed SIFs for mode I (red) and II (blue).

In this test case, the proposed method provides results within 5 % for mode I on any of the meshes. The results of mode II are a little bit worse, however the achieved error is also limited to 10 % on the coarse mesh and improves upon refinement.

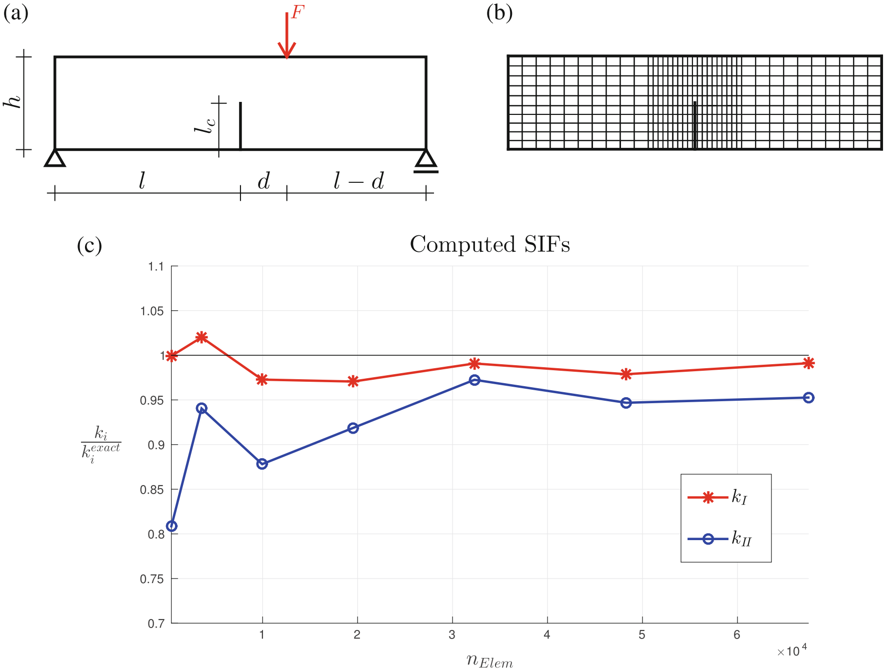


Fig. 10 Eccentric three-point bending test in two dimensions (**a**) geometry parameters, (**b**) mesh and (**c**) results

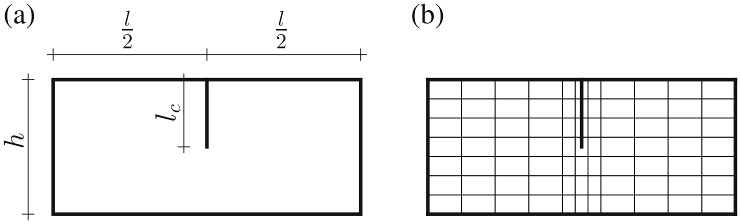


Fig. 11 Edge cracked rectangular plate (**a**) geometry parameters and (**b**) mesh

4.2 Shear Edge Crack

A shear loaded edge cracked rectangular plate is investigated next. The extent of the plate is given by: $h = 7\,\text{m}$, $l = 16\,\text{m}$ and exhibits an initial crack with a length $l_c = 3.5\,\text{m}$, as shown in Fig. 11a. SIFs are computed on 12 different meshes. We start with a coarse mesh (77 elements), which is locally refined at the crack, as illustrated in Fig. 11b. This mesh is refined until 46.025 elements are obtained.

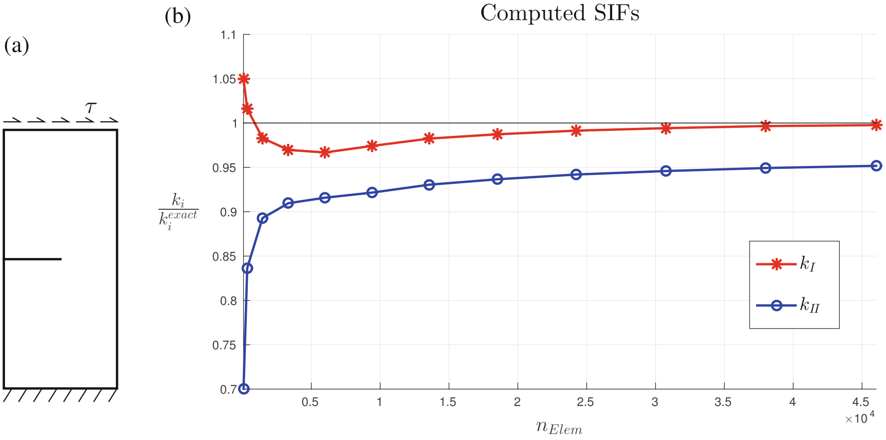


Fig. 12 Edge cracked rectangular plate (**a**) boundary and loading conditions and (**b**) results

The plate is clamped on the bottom and loaded by a shear traction $\tau = 1$ GPa on the top as Fig. 12a shows. This configuration leads to a mixed-mode loading, where the SIFs are given with $k_I = 34$ and $k_{II} = 4.55$, see e.g. [2]. The obtained results are illustrated in Fig. 12b, where the red line represents the normalized mode *I* and the blue line the normalized mode *II* SIF.

This test case also achieves good results, where the error of k_I is below 5 % for all used meshes and shows a clear convergence during the refinement. Mode *II* has again less accurate results, but also converges upon refinement.

In this two externally loaded test cases with stress-free crack paths, mode *I* SIFs were evaluated within an error of 5 % and mode *II* SIFs within 10 %. The next two cases investigate loaded crack surfaces which would lead to issues in the 'classical' interaction integral. A big advantage of the proposed method is that no modifications are necessary as the following examples show.

4.3 Loaded Crack in 2D

In this example, the accuracy of the proposed method for loaded crack surfaces is shown. Therefore, an edge cracked rectangular plate is loaded by a tension $\sigma = 1$ GPa and shear traction $\tau = 1$ GPa within the crack. The load situation and boundary conditions are illustrated in Fig. 13a. This test case uses the same geometry parameters and meshes as in Sect. 4.2. Due to the large l/h ratio it is assumed that the impact of the boundary conditions is relatively small compared to

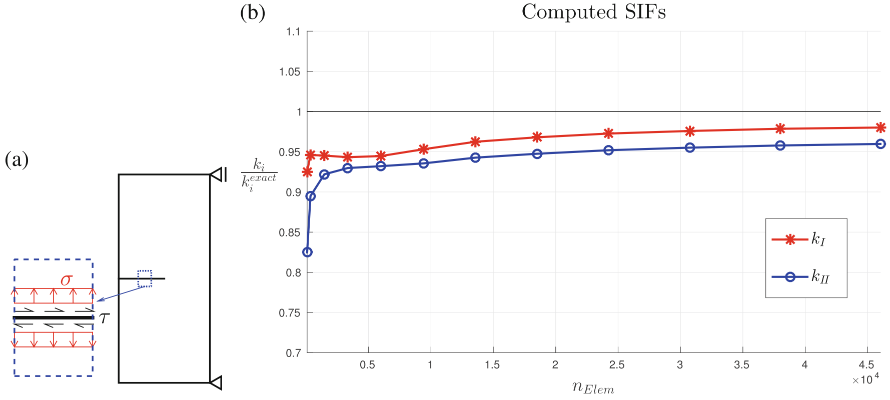


Fig. 13 Loaded crack in two dimensions (**a**) boundary and loading conditions and (**b**) results

the behaviour of the crack. This assumption leads to following expected SIFs [25]:

$$\begin{bmatrix} k_I \\ k_{II} \end{bmatrix} = \begin{bmatrix} \sigma\sqrt{\pi l_c}\left(1.122 - 0.231\frac{l_c}{h} + 10.550\frac{l_c^2}{h^2} - 21.710\frac{l_c^3}{h^3} + 30.382\frac{l_c^4}{h^4}\right) \\ \tau\sqrt{\frac{\pi l_c}{1-\frac{l_c}{h}}}\left(1.122 + 0.561\frac{l_c}{h} + 0.085\frac{l_c^2}{h^2} + 0.180\frac{l_c^3}{h^3}\right) \end{bmatrix}. \tag{24}$$

By inserting the given geometry parameters into Eq. (24), the SIFs are given by $k_I = 9.381$ and $k_{II} = 4.151$. The ratio of the computed to the expected SIFs are shown in Fig. 13b, where the red line represents mode I and the blue line mode II.

Figure 13b shows that the proposed method leads for most of the meshes to SIFs within an error of 5 %. That is, the proposed method allows the computation of SIFs also for loaded crack surfaces with a similar accuracy as for stress-free crack surfaces.

4.4 Penny-Shaped Crack

Another test case without crack propagation is considered in three dimensions. An embedded penny-shaped crack with tension and shear loaded crack surfaces is examined. The individual components of the loading are given in the global coordinate system (see Fig. 14b) by $\sigma = 1\,\text{GPa}$, $\tau_x = \tau_y = \frac{1}{\sqrt{2}}\,\text{GPa}$. The crack surface has a diameter $d = 2\,\text{m}$ and is explicitly described by 1.536 flat triangles and the crack front by 64 line segments as shown in Fig. 14a. As shown in Fig. 14c, the crack is located within a cube-like domain with a side length of 4 m. Displacements are prescribed to zero at some corner nodes on the bottom. An illustration of the situation is presented in Fig. 14. For this test case, four different meshes with

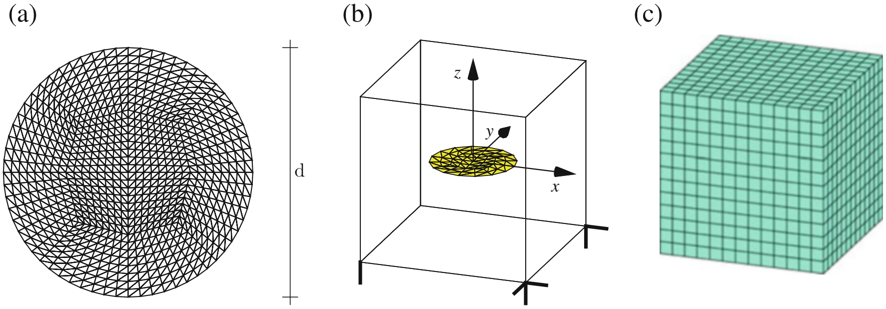


Fig. 14 Penny shaped crack (**a**) explicitly defined crack surface, (**b**) situation and (**c**) mesh

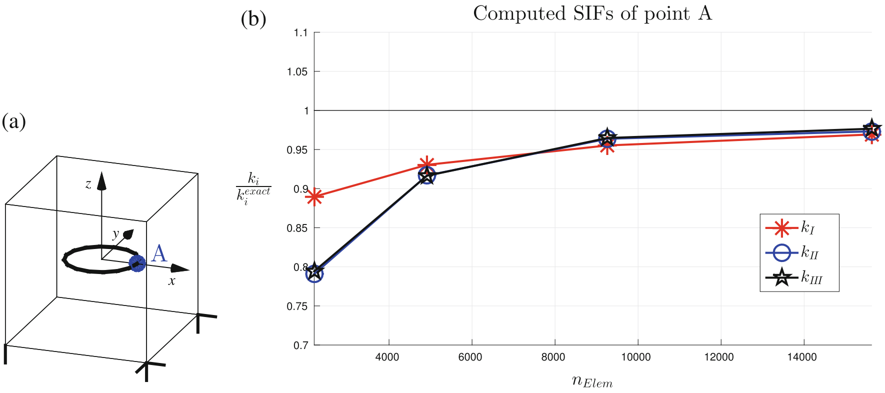


Fig. 15 Penny shaped crack (**a**) investigated point A and (**b**) results

trilinear hexahedral elements are used. The number of elements along an edge varies between 13 and 25. Figure 14c shows an example mesh with 2.197 elements.

In this general example, SIFs vary along the crack front. The expected SIFs for the whole front are given [25] by

$$\begin{bmatrix} k_I \\ k_{II} \\ k_{III} \end{bmatrix} = \begin{bmatrix} \frac{2}{\pi}\sigma\sqrt{\pi\frac{d}{2}} \\ \frac{4}{\pi(2-\nu)}\left(\tau cos\omega\right)\sqrt{\pi\frac{d}{2}} \\ \frac{4(1-\nu)}{\pi(2-\nu)}\left(\tau sin\omega\right)\sqrt{\pi\frac{d}{2}} \end{bmatrix}, \tag{25}$$

where ω describes the angle between the direction of the resultant shear traction and the reviewed point. Here, the point A (see Fig. 15a) with an angle of $\omega = \frac{\pi}{4}$ is observed and all three modes are present there. For that point, Eq. (25) leads to the following SIFs: $k_I = 11.28$, $k_{II} = 9.39$ and $k_{III} = 6.57$. The normalized computed SIFs of point A are presented in Fig. 15b.

In this example, mode *II* and mode *III* are computed with a similar accuracy as shown in Fig. 15b. Furthermore, the result shows that with only 17^3 elements, SIFs can be computed with an error of less than 10 % and 21^3 elements lead to results below 5 %. That is, SIFs are well obtained for a general three-dimensional crack configuration with stresses on the crack surface without any modifications.

4.5 Crack Propagation in Two Dimensions

Until now, only cases with straight or flat crack geometries have been investigated. To show the accuracy of the proposed method also for curved cracks, this final test case shows a mixed-mode crack propagation in a square specimen with the extent $l = 1$ m [3, 10], where curved cracks are generally expected. The displacements are prescribed on the upper and lower side of the domain with 1 mm in direction of the specified angle α, as illustrated in Fig. 16a. These boundary conditions produce an opening of the crack, wherefore no other loadings are needed. The domain is discretized by a structured mesh. It is noted that our interest is only in the computation of SIFs, wherefore no attention is given to the crack velocities. It is assumed that the crack always propagates by the crack increment da in direction of the maximum circumferential stresses. By using SIFs and the maximum circumferential stress criterion, the propagation angle θ_c can be expressed as follows [17]

$$\theta_c = 2arctan\frac{1}{4}\left(\frac{k_I}{k_{II}} \pm \sqrt{\left(\frac{k_I}{k_{II}}\right)^2 + 8}\right), \tag{26}$$

where the sign depends on the sign of k_{II}.

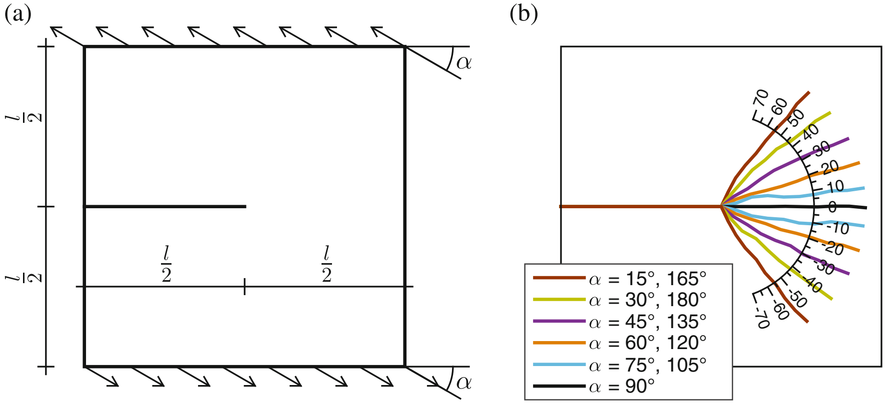


Fig. 16 Edge crack in a squared plate (**a**) geometry parameters and supports and (**b**) results of the crack propagation

This test case starts with an initial crack of length $l_c = 0.5\,\text{m}$ which is described by one straight line segment. Then, the boundary conditions are prescribed and the resulting SIFs are computed by the proposed method. With these SIFs the propagation angle θ_c is evaluated and a new line segment of the length $da = 5\,\text{cm}$ in direction of θ_c is added. This process is performed ten times. To keep the influence of the element size to a minimum, a fine setting with 101×101 elements is used. The prescribed direction of the boundary conditions is varied from $15 \leq \alpha \leq 165$. In Fig. 16b, the results of the computed crack paths are presented.

As expected, for $\alpha = 90$ the crack propagates horizontally as for a pure tension loading where only mode *I* is relevant. The greater the boundary conditions deviate from a pure tension loading, the more dominant the impact of mode *II* is. The obtained crack paths for the different α are in good agreement to [3] so that it is concluded that the proposed technique is also working for curved cracks.

5 Conclusions

In the presented work, it is shown how crack opening displacements (CODs) are used for the computation of SIFs within the XFEM in two and three dimensions. The use of CODs for computing SIFs provides an intuitive, computationally cheap and robust method, which works in two and three dimensions in a consistent manner no matter whether the crack surfaces are stress-free or loaded. The aim is to use this approach also in simulations of hydraulic fracturing where stresses are exerted by a fluid on the crack surface.

Two coordinate systems are used to simplify the definition of the enrichment functions on the basis of level-set functions and describe a reference state, where pure mode *I*, *II* and *III* openings are evaluated. These coordinate systems consider curved cracks as well as planar ones. A comparison of the approximated states and the reference states leads to the SIFs. An assignment of the fitting points to the explicitly defined crack front nodes provides SIFs for each of those nodes. At the crack tip/front, it is recommended to have rather small elements which considerably improves the solutions in general. The accuracy and robustness of the proposed method has been shown by five examples in two and three dimensions. For each test case, results were achieved below 10 % error improving upon mesh refinement. All obtained results are in good agreement with the available analytical, numerical and experimental solutions so that it is concluded that the proposed method provides good results for any crack configurations.

References

1. Anderson, T.L.: Fracture mechanics: fundamentals and applications. CRC, Boca Raton (2005)
2. Belytschko, T., Black, T.: Elastic crack growth in finite elements with minimal remeshing. Int. J. Numer. Methods Eng. **45**, 601–620 (1999)
3. Bourdin, B., Francfort, G., Marigo, J.: Numerical experiments in revisited brittle fracture. J. Mech. Phys. Solids **48**, 797–826 (2000)

4. Cazes, F., Moës, N.: Comparison of a phase-field model and of a thick level set model for brittle and quasi-brittle fracture. Int. J. Numer. Methods Eng. **103**, 114–143 (2015)
5. Chan, S.K., Tuba, I.S., Wilson, W.K.: On the finite element method in linear fracture mechanics. Eng. Fract. Mech. **2**, 1–17 (1970)
6. Dolbow, J., Gosz, M.: On the computation of mixed-mode stress intensity factors in functionally graded materials. J. Solids Struct. **39**, 2557–2574 (2002)
7. Duflot, M.: A study of the representation of cracks with level sets. Int. J. Numer. Methods Eng. **70**, 1261–1302 (2007)
8. Fett, T.: Stress Intensity Factors and Weight Functions for Special Crack Problems, vol. 6025. FZKA, Karlsruhe (1998)
9. Fries, T.: A corrected XFEM approximation without problems in blending elements. Int. J. Numer. Methods Eng. **75**, 503–532 (2008)
10. Fries, T., Baydoun, M.: Crack propagation with the extended finite element method and a hybrid explicit-implicit crack description. Int. J. Numer. Methods Eng. **89**, 1527–1558 (2012)
11. Fries, T., Belytschko, T.: The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Methods Eng. **84**, 253–304 (2010)
12. Gravouil, A., Moës, N., Belytschko, T.: Non-planar 3D crack growth by the extended finite element and level sets - Part II: level set update. Int. J. Numer. Methods Eng. **53**, 2569–2586 (2002)
13. Gray, L., Phan, A., Paulino, G., Kaplan, T.: Improved quarter-point crack tip element. Eng. Fract. Mech. **70**, 269–283 (2003)
14. Haboussa, D., Gregoire, D., Elguedj, T., Maigre, H., Combescure, A.: X-FEM analysis of the effects of holes or other cracks on dynamic crack propagations. Int. J. Numer. Methods Eng. **86**, 618–636 (2011)
15. Miehe, C., Hofacker, M., Welschinger, F.: A phase field model for rate-independent crack propagation: robust algorithmic implementation based on operator splits. Comput. Methods Appl. Mech. Eng.. **199**, 2765–2778 (2010)
16. Moës, N., Belytschko, T.: Extended finite element method for cohesive crack growth. Eng. Fract. Mech. **69**, 813–833 (2002)
17. Moës, N., Dolbow, J., Belytschko, T.: A finite element method for crack growth without remeshing. Int. J. Numer. Methods Eng. **46**, 131–150 (1999)
18. Moës, N., Gravouil, A., Belytschko, T.: Non-planar 3D crack growth by the extended finite element and level sets - Part I: mechanical model. Int. J. Numer. Methods Eng. **53**, 2549–2568 (2002)
19. Moës, N., Stolz, C., Bernard, P.E., Chevaugeon, N.: A level set based model for damage growth: the thick level set approach. Int. J. Numer. Methods Eng. **86**, 358–380 (2011)
20. Nejati, M., Paluszny, A., Zimmerman, R.: On the use of quarter-point tetrahedral finite elements in linear elastic fracture mechanics. Eng. Fract. Mech. **144**, 194–221 (2015)
21. Nikishkov, G.P.: Accuracy of quarter-point element in modeling crack-tip fields. Comput. Model. Eng. Sci. **93**, 335–361 (2013)
22. Peng, D., Merriman, B., Osher, S., Zhao, H., Kang, M.: A PDE-based fast local level set method. J. Comput. Phys. **155**, 410–438 (1999)
23. Stolarska, M., Chopp, D., Moës, N., Belytschko, T.: Modelling crack growth by level sets in the extended finite element method. Int. J. Numer. Methods Eng. **51**, 943–960 (2001)
24. Sukumar, N., Moës, N., Moran, B., Belytschko, T.: Extended finite element method for three-dimensional crack modelling. Int. J. Numer. Methods Eng. **48**, 1549–1570 (2000)
25. Tada, H., Paris, P.C., Irwin, G.R.: The Analysis of Cracks Handbook. ASME Press, New York (2000)
26. Walters, M., Paulino, G., Dodds, R.: Interaction integral procedures for 3-D curved cracks including surface tractions. Eng. Fract. Mech. **72**, 1635–1663 (2005)

Part III
Polygonal and Polyhedral Methods

The Virtual Element Method for Underground Flow Simulations in Fractured Media

Matías Fernando Benedetto, Stefano Berrone, and Andrea Borio

Abstract We focus on the problem of performing underground flow simulations in fractured media. The medium is modelled by means of the so-called Discrete Fracture Network (DFN) model. Within this framework, we discuss about the use of the Virtual Element Method (VEM) in performing simulations, and about its role in facilitating the meshing process. DFN models are characterized by a large number of planar fractures stochastically generated and placed in a 3D reference system starting from given probabilistic distribution of space position, dimension, aspect ratio, orientation and hydrogeological properties. For these reasons, realistic DFNs are usually characterized by the presence of very complex geometrical configurations on which it is very difficult to obtain a good quality Finite Element mesh. In particular, in order to circumvent these difficulties, we consider two different VEM approaches: an approach based on a totally conforming polygonal mesh obtained starting from a triangular mesh independently generated on the different fractures of the system and an approach based on a partially conforming mesh in which the connection between the fracture meshes is imposed through a mortaring approach borrowed from the classical domain decomposition methods. Concerning these two approaches, we discuss their relative advantages and present numerical results aimed at comparing them.

1 Introduction

The focus of the present contribution is on underground flow simulations in fractured media. This is a quite relevant issue in several present critical applications, concerned with the long term analysis of geological systems subject to underground exploitation.

M.F. Benedetto
Laboratorio de Métodos Numéricos en Ingeniería, Facultad de Ingeniería,
Universidad de Buenos Aires, Av. Las Heras 2214, C1127AAR, Buenos Aires, Argentina
e-mail: mbenedetto@fi.uba.ar

S. Berrone (✉) • A. Borio
Dipartimento di Scienze Matematiche, Politecnico di Torino,
Corso Duca degli Abruzzi, 24, 10129 Torino, Italy
e-mail: stefano.berrone@polito.it; andrea.borio@polito.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_8

A possible representation of a fractured medium is given by the so-called Discrete Fracture Network (DFN) model [1, 23, 26]. Assuming that the surrounding rock matrix is impervious, according to DFN model, the fluid flows along fractures and through fractures intersections. Most methods proposed in the recent literature concerning flow simulations in fracture networks ask for some kind of mesh conformity at fracture intersections (traces), in order to ensure a correct representation of the flux exchange between fractures. This (possibly partial) conformity requirement often results in a fully coupled meshing process [25, 30, 36–38, 41], which is quite often very difficult to be performed or even infeasible, in some critical configurations, which are nevertheless likely to be encountered, especially in the framework of uncertainty quantification analysis [18]. To remove these conformity constraints, an optimization based method was proposed in [14–16, 20] and developed in [9, 17, 19]. More details on this optimization approach can be found in [39]. A different approach not involving mesh generation is described in [24, 34, 35]. More complex models on the fractures can be found in [27–29, 31].

In the present work, we discuss on the use of the rather new Virtual Element Method for performing simulations in fracture networks. After introducing some notation in the sequel of this section, we will describe the problem formulation in Sect. 2; in Sect. 3 we will discuss the use of VEM in this framework, focusing on two possible approaches. Finally, we propose in Sect. 4 some numerical results aimed at comparing the two approaches.

1.1 Discrete Fracture Networks

Each fracture in the underground is represented as a (2D) open polygon. In what follows, we will identify the geological fracture and its representing polygon, and we will refer also to the polygons as fractures. Accordingly, a DFN Ω is represented as a (3D) set given by the union of N fractures F_i, with $i \in \mathcal{I} = \{1, \ldots, N\}$. We also introduce the set $\partial\Omega = \cup_{i\in\mathcal{I}}\partial F_i$. Fractures intersections are called *traces* and denoted by Γ_m, with $m \in \mathcal{M} = \{1, \ldots, M\}$ (see Fig. 1). Without loss of generality, we assume that the set $\bar{\Omega}$ is connected. For the sake of simplicity, we also assume that each trace is given by the intersection of precisely two fractures. For future reference, we introduce the following notation:

- $\forall i \in \mathcal{I}$, we introduce the subset $\mathcal{M}_i \subset \mathcal{M}$ of the indices of traces lying on F_i; each subset $\mathcal{M}_i$ is assumed to be ordered, and we will denote by $\mathcal{M}_i(k)$ the k-th index of a trace in $\mathcal{M}_i$;
- $\forall m \in \mathcal{M}$, we introduce the couple $\mathcal{I}_m = (i, j)$ of indices such that $\Gamma_m = F_i \cap F_j$; the couple is assumed to be ordered in such a way that $i \leq j$;
- for each $i \in \mathcal{I}$ and each $m \in \mathcal{M}_i$, we fix a unit vector $\hat{n}_m^i$ normal to Γ_m on F_i.

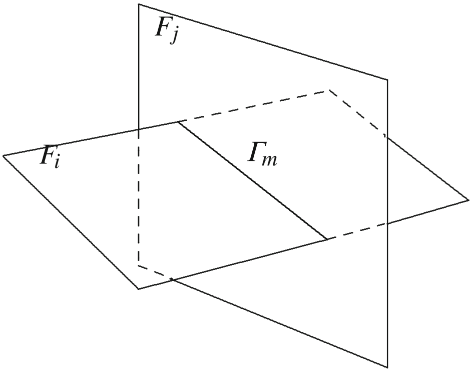


Fig. 1 Example of two fractures (F_i and F_j) intersecting and generating a trace (Γ_m)

1.2 Further Notation

Given an arbitrary set ω, we denote by $(\cdot,\cdot)_\omega$ and $\|\cdot\|_\omega$ the $\mathrm{L}^2(\omega)$ scalar product and norm, respectively; furthermore, we let $(\cdot,\cdot)_{\alpha,\omega}$ and $\|\cdot\|_{\alpha,\omega}$ denote the $\mathrm{H}^\alpha(\omega)$ scalar product and norm, respectively. For any function defined on Ω, we use a subscript i to indicate its restriction to fracture F_i. Let $\Gamma^D \subseteq \partial\Omega$ be the portion of boundary on which Dirichlet conditions are imposed, and let h^D be the Dirichlet condition defined on Γ^D. We define the functional spaces

$$\begin{aligned} V_i &:= \left\{v \in \mathrm{H}^1(F_i) : \gamma_{\Gamma^D}(v) = 0\right\} && \forall i \in \mathcal{I}\,, \\ V_i^D &:= \left\{v \in \mathrm{H}^1(F_i) : \gamma_{\Gamma^D}(v) = h_i^D\right\} && \forall i \in \mathcal{I}\,, \\ V &:= \{v : v_i \in V_i \quad \forall i \in \mathcal{I}\}\,, \\ V^D &:= \left\{v : v_i \in V_i^D \quad \forall i \in \mathcal{I}\right\}. \end{aligned}$$

For any segment $\sigma \subset F_i$, $i \in \mathcal{I}$, we introduce the trace operator $\gamma_\sigma : \mathrm{H}^1(F_i) \to \mathrm{H}^{\frac{1}{2}}(\sigma)$ and the notation

$$\langle \mu, \beta \rangle_\sigma := {}_{\mathrm{H}^{-\frac{1}{2}}(\sigma)}\langle \mu, \beta \rangle_{\mathrm{H}^{\frac{1}{2}}(\sigma)}\,, \qquad \forall \mu \in \mathrm{H}^{-\frac{1}{2}}(\sigma)\,,\ \beta \in \mathrm{H}^{\frac{1}{2}}(\sigma)\,,$$

to denote the duality product between $\mathrm{H}^{-\frac{1}{2}}(\sigma)$ and $\mathrm{H}^{\frac{1}{2}}(\sigma)$. Let $v \in V$. In order to simplify the notation, it is convenient to introduce the vectors $\gamma_{\mathcal{M}_i}(v)$, $\forall i \in \mathcal{I}$, the k-th element of $\gamma_{\mathcal{M}_i}(v)$ being $\gamma_{\Gamma_{\mathcal{M}_i(k)}}(v)$. Furthermore, we introduce the jump across a trace Γ_m as

$$[\![v]\!]_{\Gamma_m} := \gamma_{\Gamma_m}(v_i) - \gamma_{\Gamma_m}(v_j)\,, \quad \text{if } \mathcal{I}_m = (i,j)\,,$$

and we introduce the symbols $[\![v]\!]_{\mathcal{M}}$ and $[\![v]\!]_{\mathcal{M}_i}$ to denote the vectors of jumps of v across all traces in the network, and across traces on F_i, respectively. With the same purpose, $\forall i \in \mathcal{I}$, we introduce the notation

$$\langle \mu, \beta \rangle_{\mathcal{M}_i} := \sum_{m \in \mathcal{M}_i} \langle \mu_m, \beta_m \rangle_{\Gamma_m}, \ \forall \mu \in \prod_{m \in \mathcal{M}_i} \mathrm{H}^{-\frac{1}{2}}(\Gamma_m), \beta \in \prod_{m \in \mathcal{M}_i} \mathrm{H}^{\frac{1}{2}}(\Gamma_m) \ .$$

Finally, for any $\beta, \lambda \in \prod_{m \in \mathcal{M}_i} \mathrm{H}^{\frac{1}{2}}(\Gamma_m)$ we denote

$$(\beta, \lambda)_{\mathcal{M}_i} := \sum_{m \in \mathcal{M}_i} (\beta, \lambda)_{\Gamma_m} \ .$$

2 Problem Formulation

We are interested in computing the hydraulic head $h = \pi/(\rho g) + z \in V^D$, where π is the fluid pressure, g the gravitational acceleration, ρ the fluid density and z the elevation. The hydraulic head, on each fracture F_i, is modeled by means of the Darcy law as follows. Let K_i denote the transmissivity on F_i, which we assume to be constant, and $f_i = f_i(x)$ denote the source term on F_i; notice that both K_i and f_i are functions of the local tangential coordinate system.

The problem on each fracture is: find $h_i \in V_i^D$ such that, $\forall v_i \in V_i$,

$$(\mathrm{K}_i \nabla h_i, \nabla v_i)_{F_i} = (f_i, v_i)_{F_i} + \left\langle h_i^N, \gamma^i_{\Gamma_i^N}(v_i) \right\rangle_{\Gamma_i^N} + \left\langle \left[\!\!\left[\frac{\partial h_i}{\partial \hat{n}_{\mathcal{M}_i}} \right]\!\!\right]_{\mathcal{M}_i}, \gamma_{\mathcal{M}_i}(v_i) \right\rangle_{\mathcal{M}_i} \tag{1}$$

where $\left[\!\!\left[\frac{\partial h_i}{\partial \hat{n}^i_m} \right]\!\!\right]_{\Gamma_m}$ is the jump of the co-normal derivative $\frac{\partial h_i}{\partial \hat{n}^i_m} = \mathrm{K}_i \nabla h_i \cdot \hat{n}^i_m$ along $\hat{n}^i_m$; furthermore, $\Gamma_i^N \subseteq \partial F_i$ is the Neumann boundary on F_i and $h_i^N \in \mathrm{H}^{-\frac{1}{2}}\left(\Gamma_i^N\right)$ is the Neumann boundary condition. For future reference, we set $\Gamma^N = \cup_{i \in \mathcal{I}} \Gamma_i^N \subset \partial\Omega$ and define h^N such that h_i^N is the restriction of h^N to F_i.

The problems on each fracture are coupled by the imposition of suitable matching conditions which guarantee the continuity of the solution and balance of incoming and outgoing fluxes at each trace: $\forall m \in \mathcal{M}$, with $\mathcal{I}_m = (i,j)$, we have

$$\begin{aligned} [\![h]\!]_{\Gamma_m} &= \gamma_{\Gamma_m}(h_i) - \gamma_{\Gamma_m}(h_j) = 0, \\ \left[\!\!\left[\frac{\partial h_i}{\partial \hat{n}^i_m} \right]\!\!\right]_{\Gamma_m} &+ \left[\!\!\left[\frac{\partial h_j}{\partial \hat{n}^j_m} \right]\!\!\right]_{\Gamma_m} = 0. \end{aligned} \tag{2}$$

3 The Virtual Element Method for DFN Simulations

First introduced in [3] and extended in [2, 4–7, 22], the Virtual Element Method allows the use of any kind of non-degenerate star-shaped polygon to mesh the spatial domain, even including the possibility of straight angles. In the present framework, we take advantage from this flexibility to easily build a mesh which, on each fracture, is locally conforming to the traces. In the following of this section, we review the use of VEM, focusing on the framework of DFN simulations.

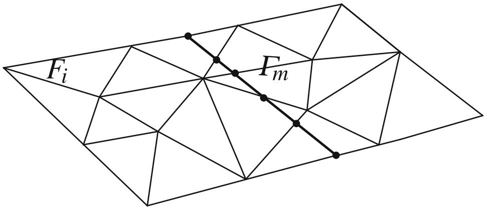
3.1 The VEM Setting in the DFN Framework

Let us fix a fracture F_i. To obtain a locally conforming mesh, we first introduce on F_i a triangular mesh built independently of trace positions; the triangles are then cut into polygons by the traces, possibly prolonging the trace segment up to the nearest mesh edge if it happens to end in the interior of a triangle. Note that in this latter case the trace tip is kept as a node of the discretization, a new node is created at the intersection between the prolongation of the trace and the mesh edge, and therefore two edges are created, with a 180° angle between them. Let $\mathcal{T}_{\delta i}$ be the resulting local mesh. Referring to the configuration already depicted in Fig. 1, we report in Figs. 2 and 3 an example of the VEM meshes $\mathcal{T}_{\delta i}$ and $\mathcal{T}_{\delta j}$ obtained on the two fractures F_i and F_j, respectively. Furthermore, we notice that all polygons created with the above procedure are convex, thus satisfying the assumptions in [3], that require the elements to be star-shaped with respect to a ball. Let $\mathcal{T}_\delta = \cup_{i\in\mathcal{I}}\mathcal{T}_{\delta i}$, being δ a global mesh parameter, e.g. the maximum element diameter. We will use the symbols $\mathcal{E}_{\delta i}$ and $\mathcal{V}_{\delta i}$ to denote the sets of edges and vertices on fracture F_i, respectively; similarly, $\mathcal{E}_{\delta i}^E$ and $\mathcal{V}_{\delta i}^{\mathrm{E}}$ denote sets of edges and vertices on the element $E \in \mathcal{T}_{\delta i}$, and define $\mathcal{E}_\delta = \cup_{i\in\mathcal{I}}\mathcal{E}_{\delta i}$, $\mathcal{V}_\delta = \cup_{i\in\mathcal{I}}\mathcal{V}_{\delta i}$.

Let $k \in \mathbb{N}$ be the VEM order, let $i \in \mathcal{I}$ and consider an element $E \in \mathcal{T}_{\delta i}$. In order to define the Virtual Element space of order k for the DFN, let us define the local finite dimensional functional space

$$V_{\delta i}^E := \left\{ v \in \mathrm{H}^1(E) : \gamma_e(v) \in \mathbb{P}^k(e) \ \forall e \subset \partial E,\ \Delta v|_E \in \mathbb{P}^{k-2}(E)\,, \right\}\,, \tag{3}$$

Fig. 2 VEM mesh $\mathcal{T}_{\delta i}$ on F_i



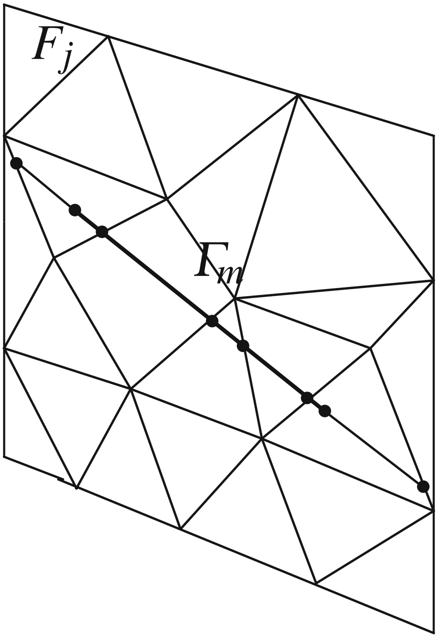


Fig. 3 VEM mesh $\mathcal{T}_{\delta j}$ on F_j

and let the virtual elements space on F_i be

$$V_{\delta i} := \left\{ v_\delta \in \mathbb{C}^0\left(F_i\right) : v \in V_{\delta i}^E \quad \forall E \in \mathcal{T}_{\delta i} \right\} .$$

Finally, we introduce the following global discrete subspace of V:

$$V_\delta := \{ v_\delta \in V : v \in V_{\delta i} \quad \forall i \in \mathcal{I} \} . \tag{4}$$

It can be proved (see [3]) that a possible set of degrees of freedom that uniquely define a function $v_\delta \in V_\delta$ is given by:

- the values of v_δ at each vertex $V \in \mathcal{V}_\delta$;
- if $k > 1$, the values of v_δ at $k - 1$ internal points (e.g. internal Gauss-Lobatto quadrature nodes) on each edge $e \in \mathcal{E}_\delta$;
- if $k > 1$ the moments $\frac{1}{|E|} \int_E v_h m_{\boldsymbol{\alpha}}$ for $| \boldsymbol{\alpha} | \leq k - 2$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{N}^2$ and

$$m_{\boldsymbol{\alpha}}(x, y) := \left(\frac{x - x_E}{h_E} \right)^{\alpha_1} \left(\frac{y - y_E}{h_E} \right)^{\alpha_2} ,$$

where (x_E, y_E) and h_E are the centroid and the diameter of the element E, respectively. As a basis of V_δ, we consider the functions ϕ_k, $k \in \{1, \ldots, N\}$ which are lagrangian with respect to the above degrees of freedom.

We now turn our attention to Eq. (1), assuming that the transmissivity coefficient is constant on each element of the mesh. Following [3, 6], in order to discretize this equation using functions in V_δ, since the virtual functions are not known in the interior of the elements, we introduce, for each $i \in \mathcal{I}$ and $E \in \mathcal{T}_{\delta i}$, the operator $\Pi^\nabla_E \colon V^E_{\delta i} \to \mathbb{P}^k(E)$ that associates, to each $\phi \in V^E_{\delta i}$, the polynomial $\Pi^\nabla_E \phi$ such that

$$\begin{cases} \left(\mathrm{K}_i \nabla \Pi^\nabla_E \phi, \nabla p\right)_E = (\mathrm{K}_i \nabla \phi, \nabla p)_E & \forall p \in \mathbb{P}^k(E)\,, \\ \sum_{V \in \mathcal{V}^{\mathrm{E}}_{\delta i}} \Pi^\nabla_E \phi(V) = \sum_{V \in \mathcal{V}^{\mathrm{E}}_{\delta i}} \phi(V) & \text{if } k = 1, \\ \int_E \Pi^\nabla_{E,k} v_\delta = \int_E v_\delta & \text{if } k > 1. \end{cases}$$

We remark that the application of this operator only requires the knowledge of the degrees of freedom of ϕ, through the application of Green's formula (see [5]). In addition to Π^∇_E, for each $E \in \mathcal{T}_{\delta i}$, let $S^E \colon V^E_{\delta i} \times V^E_{\delta i} \to \mathbb{R}$ be a symmetric bilinear form defined in such a way that there exist two positive constants c_* and c^* independent on E and i such that $\forall \phi \in V^E_{\delta i}$, if $\Pi^\nabla_E \phi = 0$, then

$$c_* \left(\mathrm{K}_i \nabla \phi, \nabla \phi\right)_E \leq S^E\left(\phi, \phi\right) \leq c^* \left(\mathrm{K}_i \nabla \phi, \nabla \phi\right)_E . \tag{5}$$

With the above ingredients, we define the discrete bilinear form $a^E_\delta \colon V_\delta \times V_\delta \to \mathbb{R}$ such that, $\forall v, w \in V_\delta$,

$$a^E_\delta\left(v, w\right) := \left(\mathrm{K}_i \nabla \Pi^\nabla_E v, \nabla \Pi^\nabla_E w\right)_E + S^E\left(v - \Pi^\nabla_E v, w - \Pi^\nabla_E w\right) .$$

Thanks to (5), we easily find that $a^E_\delta\left(v, v\right)$ scales like $(\mathrm{K}_i \nabla v, \nabla v)_E$, with scaling constants independent of E and of the fracture index i. Thus, if we introduce the fracture-wide bilinear form $a_{\delta i} \colon V_{\delta i} \times V_{\delta i} \to \mathbb{R}$ such that

$$a_{\delta i}\left(v, w\right) := \sum_{E \in \mathcal{T}_{\delta i}} a^E_\delta\left(v, w\right) ,$$

we have the following property:

$$\exists \alpha_*, \alpha^* > 0 \colon \alpha_* \left(\mathrm{K}_i \nabla v, \nabla v\right)_{F_i} \leq a_{\delta i}\left(v, v\right) \leq \alpha^* \left(\mathrm{K}_i \nabla v, \nabla v\right)_{F_i} . \tag{6}$$

Moreover, since the scalar product (f, v_δ) is not computable in general if v_δ is a virtual function, we define the discrete scalar product

$$(f_i, v_{\delta i})_{\delta, F_i} := \left(f_i, \tilde{\Pi}^0_k v_{\delta i}\right)_{F_i} \quad \forall i \in \mathcal{I} ,$$

where the pseudo-projection $\tilde{\Pi}^0_k$ is defined, as done in [5], by local projections, recovering from $\Pi^\nabla_{E,k} v_{\delta i}$ the missing information about $v_{\delta i}$:

$$\forall E \in \mathcal{T}_{\delta i}, \begin{cases} \left(\tilde{\Pi}^0_k v_{\delta i}, p\right)_E = (v_{\delta i}, p)_E & \forall p \in \mathbb{P}^{k-2}(E) , \\ \left(\tilde{\Pi}^0_k v_{\delta i}, p\right)_E = \left(\Pi^\nabla_{E,k} v_{\delta i}, p\right)_E & \forall p \in \mathbb{P}^k(E) \setminus \mathbb{P}^{k-2}(E) . \end{cases}$$

Finally, to ease the notation in the following it is convenient to define the global discrete products

$$a_\delta(v,w) := \sum_{i\in\mathcal{I}} a_{\delta i}(v,w) \qquad \forall v,w \in V_\delta,$$

$$(f,v_\delta)_\delta := \sum_{i\in\mathcal{I}} (f_i, v_{\delta i})_{\delta,F_i} \qquad \forall v \in V_\delta.$$

Remark 1 As suggested in [3, 6], a possible choice for the stabilization term S^E is given by the scalar product between the vectors containing the degrees of freedom of the two arguments on the element. This choice guarantees property (5) under some basic regularity assumptions on the triangulation. This choice has been adopted in [9].

3.2 Formulation of the Problem Towards Domain Decomposition

For each $m \in \mathcal{M}$, with $\mathcal{I}_m = (i,j)$, we define the function $s_{\Gamma_m}: \mathcal{I}_m \to \{0,1\}$ such that

$$s_{\Gamma_m}(i) = 1, \quad s_{\Gamma_m}(j) = 0,$$

and the bilinear form $b_i: M_i := \prod_{m\in\mathcal{M}_i} \mathrm{H}^{-\frac{1}{2}}(\Gamma_m) \times V_i \to \mathbb{R}$ such that

$$b_i(v,\psi) := \sum_{m\in\mathcal{M}_i} (-1)^{s_{\Gamma_m}(i)} \langle \psi_m, \gamma_{\Gamma_m}(v_i)\rangle_{\Gamma_m},$$

in such a way that

$$b(v,\psi) := \sum_{i\in\mathcal{I}} b_i(v,\psi) = \sum_{m\in\mathcal{M}} \left\langle \psi_m, [\![v]\!]_{\Gamma_m}\right\rangle_{\Gamma_m}.$$

We define the functional $\mathcal{F}: V \times M := \prod_{m\in\mathcal{M}} \mathrm{H}^{-\frac{1}{2}}(\Gamma_m)$ such that

$$\mathcal{F}(v,\psi) := \sum_{i\in\mathcal{I}} \frac{1}{2}(\mathrm{K}_i \nabla v_i, \nabla v_i)_{F_i} - (f_i, v_i)_{F_i} - \left\langle h_i^N, \gamma^i_{\Gamma_i^N}(v_i)\right\rangle_{\Gamma_i^N}$$
$$+ \left(\nabla \mathcal{R}_i\left(h_i^D\right), \nabla v_i\right)_{F_i} + b_i(v,\psi) + b_i\left(\mathcal{R}_i\left(h^D\right), \psi\right),$$

where $\mathcal{R}_i$ is the lift operator from $\mathrm{H}^{\frac{1}{2}}(F_i)$ to $\mathrm{H}^1(F_i)$, $i \in \mathcal{I}$. It is well known (see [40]) that solving problem (1)–(2) is equivalent to solve the problem of finding $(h, \lambda) \in V^D \times M$ such that $h_i = h_i^0 + \mathcal{R}_i\left(h^D\right)$, $h_i^0 \in V_i$, $i \in \mathcal{I}$ and

$$\mathcal{F}\left(h^0, \lambda\right) = \min_{v \in V} \max_{\psi \in M} \mathcal{F}(v, \psi), \tag{7}$$

that, by uniqueness of the solution, implies

$$\lambda_m = \left[\!\left[\frac{\partial h_i}{\partial \hat{n}_m^i} \right]\!\right]_{\Gamma_m} = - \left[\!\left[\frac{\partial h_j}{\partial \hat{n}_m^j} \right]\!\right]_{\Gamma_m},$$

with $\mathcal{I}_m = (i,j)$. Notice that the functional $\mathcal{F}$ is made up of local contributions from each fracture. We will now present two different approaches for discretizing problem (7).

3.3 A Globally Conforming Approach

A first approach to tackle problem (7), introduced in [10], is to strongly impose the matching conditions by building a globally conforming mesh. This is easily achieved by exploiting the capability of the Virtual Element Method to handle straight angles. Consider the mesh $\mathcal{T}_\delta$ constructed as depicted in Sect. 3.1 and let us consider an arbitrary trace Γ_m, with $m \in \mathcal{M}$ and $\mathcal{I}_m = (i,j)$. Then, we add to $\mathcal{T}_{\delta i}$ the nodes generated by $\mathcal{T}_{\delta j}$ on Γ_m, and vice-versa. Some polygons belonging to mesh $\mathcal{T}_{\delta i}$ ($\mathcal{T}_{\delta j}$, respectively) having an edge lying on Γ_m, will possibly have such edges split by the new nodes, the new edges forming a straight angle at their intersection. In the same configuration of Figs. 1, 2 and 3, we show in Figs. 4 and 5 the globally conforming VEM meshes on F_i and F_j, respectively.

We will call $\mathcal{T}_\delta^{\mathrm{gc}}$ this new set of globally conforming polygons and discretize problem (7) with the Virtual Element Method as previously described. Let $V_\delta^{\mathrm{gc}} \subset V$ be the Virtual Element space defined on $\mathcal{T}_\delta^{\mathrm{gc}}$ considering, for each trace node, two different degrees of freedom, each one associated with one of the two fractures intersecting there. In other words, we admit that functions in V_δ^{gc} could be discontinuous on each trace, even though on each trace their degrees of freedom

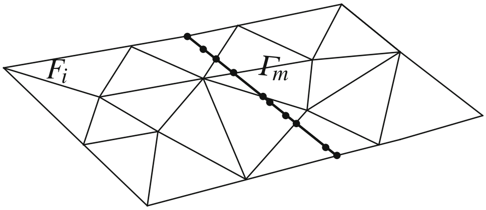


Fig. 4 Globally conforming VEM mesh on F_i

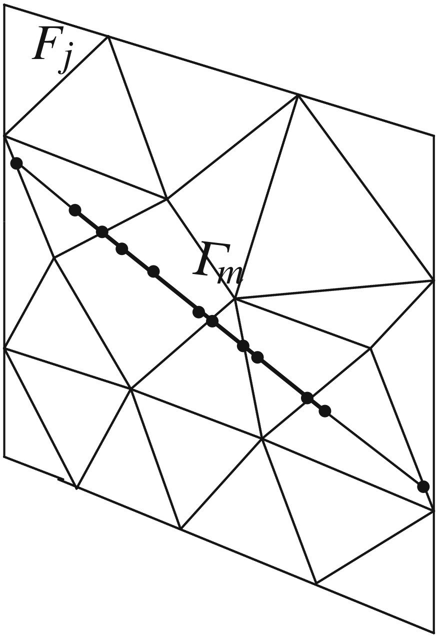


Fig. 5 Globally conforming VEM mesh on F_j

correspond to the same geometrical point. To discretize the space M we define, for each $i \in \mathcal{I}$ and on each trace Γ_m such that $m \in \mathcal{M}_i$, the finite dimensional space

$$M^{\mathrm{gc}}_{\delta m,i} = \operatorname{span}\left\{\mu^m_{ki},\ k = 1,\dots,N_{\Gamma_m}\right\}\ ,$$

where N_{Γ_m} is the number of interior nodes on Γ_m and μ^m_{ki} is a continuous linear operator such that

$$\left\langle \mu^m_{ki}, \gamma_{\Gamma_m}\left(v_{\delta j}\right)\right\rangle_{\Gamma_m} = \delta_{ij} v_{\delta i}(\mathbf{x}^m_k) \quad \forall v_\delta \in V^{\mathrm{gc}}_\delta, \tag{8}$$

being δ_{ij} the Kronecker delta and $\mathbf{x}^m_k$ the k-th node on trace Γ_m. The space that discretizes M is $M^{\mathrm{gc}}_\delta = \prod_{m\in\mathcal{M}} M^{\mathrm{gc}}_{\delta m}$, where

$$M^{\mathrm{gc}}_{\delta m} := \left\{\mu^m_k \colon \mu^m_k = \mu^m_{ki} - \mu^m_{kj} \text{ if } \mathcal{I}_m = (i,j)\ ,\ k = 1,\dots,N_{\Gamma_m}\right\}\ .$$

With this definition we have, for any $m \in \mathcal{M}$, if $\mathcal{I}_m = (i,j)$, $\forall k \in \{1,\dots,N_{\Gamma_m}\}$,

$$\left\langle \mu^m_k, [\![v_\delta]\!]_{\Gamma_m}\right\rangle_{\Gamma_m} = \left\langle \mu^m_{ki}, \gamma_{\Gamma_m}\left(v_{\delta i}\right)\right\rangle_{\Gamma_m} - \left\langle \mu^m_{kj}, \gamma_{\Gamma_m}\left(v_{\delta j}\right)\right\rangle_{\Gamma_m} = v_{\delta i}(\mathbf{x}^m_k) - v_{\delta j}(\mathbf{x}^m_k),$$

and therefore, enforcing orthogonality of the jumps on traces with respect to M_δ^{gc} is enough to obtain continuity on V_δ^{gc}. Indeed,

$$b\left(v_\delta, \psi_\delta\right) = 0 \quad \forall \psi_\delta \in M_\delta^{\mathrm{gc}} \iff \llbracket v_\delta \rrbracket_{\mathcal{M}} = 0\,.$$

Let us define, for $i \in \mathcal{I}$

$$h_\delta = h_\delta^0 + \mathcal{R}_\delta\left(h^D\right),$$

where $h_\delta^0 \in V_\delta^{\mathrm{gc}}$ and $\mathcal{R}_\delta\left(h^D\right)$ is the discrete lifting of the boundary conditions. For the sake of simplicity we neglect the approximation error of the Dirichlet boundary condition. The globally conforming discrete solution $(h_\delta^0, \lambda_\delta) \in V_\delta^{\mathrm{gc}} \times M_\delta^{\mathrm{gc}}$ satisfies the virtual element discretization of (7):

$$\mathcal{F}\left(h_\delta^0, \lambda_\delta\right) = \min_{v_\delta \in V_\delta^{\mathrm{gc}}} \max_{\psi_\delta \in M_\delta^{\mathrm{gc}}} \mathcal{F}\left(v_\delta, \psi_\delta\right)\ ,$$

which leads us to the following saddle point formulation:

$$\begin{cases} a_\delta\left(h_\delta^0, v_\delta\right) + b\left(v_\delta, \lambda_\delta\right) = \left(f, v_\delta\right)_\delta + \left\langle h^N, \gamma_{\Gamma^N}\left(v_{\delta i}\right)\right\rangle_{\Gamma^N} & \\ \qquad\qquad\qquad\qquad + a_\delta\left(\mathcal{R}_\delta\left(h^D\right), v_\delta\right) & \forall v_\delta \in V_\delta^{\mathrm{gc}}, \\ b\left(h_\delta^0, \psi_\delta\right) = -b\left(\mathcal{R}_\delta\left(h^D\right), \psi_\delta\right) & \forall \psi_\delta \in M_\delta^{\mathrm{gc}}. \end{cases}$$

The above problem has a unique solution because $v_\delta \mapsto \sum_{i \in \mathcal{I}} \left(\mathrm{K}_i \nabla v_{\delta i}, \nabla v_{\delta i}\right)_{F_i}$ is a norm on

$$\begin{aligned} W_\delta^{\mathrm{gc}} := \left\{v_\delta \in V_\delta^{\mathrm{gc}} \colon b\left(v_\delta, \psi_\delta\right) = 0 \quad \forall \psi_\delta \in M_\delta^{\mathrm{gc}}\right\} = \\ = \left\{v_\delta \in V_\delta^{\mathrm{gc}} \colon \llbracket v_\delta \rrbracket_{\Gamma_m} = 0 \quad \forall m \in \mathcal{M}\right\}\ , \end{aligned}$$

and therefore a_δ is coercive on W_δ^{gc} thanks to (6), and that

$$\forall \psi_\delta \in M_\delta, \quad \sup_{v_\delta \in V_\delta^{\mathrm{gc}}} \frac{b\left(v_\delta, \psi_\delta\right)}{\left\|v_\delta\right\|_{V_\delta}} = \left\|\psi_\delta\right\|_{M_\delta}\ ,$$

being M_δ a space of linear operators.

3.3.1 Implementation

Let $\left\{\phi_k^{\mathrm{gc}}\right\}_{k=1}^{\dim V_\delta^{\mathrm{gc}}}$ be the Lagrangian basis of V_δ^{gc} and define, for each $i \in \mathcal{I}$, the fracture stiffness matrix A_i, such that

$$(A_i)_{kl} = a_{\delta i}\left(\phi_{ki}^{\mathrm{gc}}, \phi_{li}^{\mathrm{gc}}\right) .$$

We note that, by (8), if ϕ_l^{gc} is a basis function associated to an internal node of the mesh (always placed on the boundary of a VEM element),

$$b_i\left(\phi_l^{\mathrm{gc}}, \mu_k^m\right) = \begin{cases} (-1)^{s_{\Gamma_m}(i)} & \text{if } \mathbf{x}_l = \mathbf{x}_k^m, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathbf{x}_l$ is the node associated to ϕ_l^{gc}. If ϕ_l^{gc} is a basis function associated to one of the polygon internal degrees of freedom, $b_i\left(\phi_l, \mu_k^m\right) = 0\ \forall i \in \mathcal{I}, m \in \mathcal{M}_i, k \in \{1, \ldots, N_{\Gamma_m}\}$. To collect the terms coming from the bilinear form b, we define a global numbering of the degrees of freedom on all the traces and, for each $m \in \mathcal{M}$, with $\mathcal{I}_m = (i, j)$, we identify the row vector B_m such that $(B_m)_k = 1$ if the k-th trace degree of freedom is on trace m and fracture i and $(B_m)_k = -1$ if the k-th trace degree of freedom is on trace m and fracture j. Then the vector $\mathbf{h}$ containing the degrees of freedom of h_δ is the solution of

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{d} \end{pmatrix} , \tag{9}$$

where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers, $\mathbf{f}$ is the vector containing the right-hand-side terms, $\mathbf{d}$ the vector of nodal values of h^D on the traces and

$$A := \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_N \end{pmatrix} , \qquad B := \begin{pmatrix} B_1 \\ \vdots \\ B_M \end{pmatrix} .$$

Using classical results (see e.g. [33]) it is easily proven that system (9) has a unique solution. Moreover, this reformulation falls into the framework of domain decomposition methods [43]; this property is exploited in [10] to devise a preconditioned one-level FETI method [32] for its solution.

3.4 A Hybrid Mortar Virtual Element Approach

The second approach we present here, developed in [11], consists in imposing a weak continuity of the solution, by applying the mortar element method [8, 12, 13].

We consider the locally conforming mesh $\mathcal{T}_\delta$ defined in Sect. 3.1 and, for each $m \in \mathcal{M}$, with $\mathcal{I}_m = (i,j)$, we introduce a finite dimensional space $M_{\delta m} \subset \mathrm{L}^2(\Gamma_m)$ defined on the discretization of the trace induced by $\mathcal{T}_{\delta i}$. We discretize (7) on $V_\delta \times M_\delta = \prod_{m\in\mathcal{M}} M_{\delta m}$, where V_δ is the Virtual Element space defined by (3)–(4). The mortar formulation of the problem is: find $h_\delta = h_\delta^0 + \mathcal{R}_\delta\left(h^D\right)$, with $h_\delta^0 \in V_\delta$ and $\lambda_\delta \in M_\delta$ such that,

$$\begin{cases} a_\delta\left(h_\delta^0, v_\delta\right) + b\left(v_\delta, \lambda_\delta\right) = \left(f, v_\delta\right)_\delta + \left(h^N, v_\delta\right)_{\Gamma^N} & \\ \qquad\qquad\qquad\qquad - a_\delta\left(\mathcal{R}_\delta\left(h^D\right), v_\delta\right) & \forall v_\delta \in V_\delta\,, \\ b\left(h_\delta^0, \psi_\delta\right) = -b\left(\mathcal{R}_\delta\left(h^D\right), \psi_\delta\right) & \forall \psi_\delta \in M_\delta\,. \end{cases} \tag{10}$$

In practical implementations, $M_{\delta m}$ will be a piecewise polynomial space (see [12, 42]). The well-posedness of (10) is proved in [11] under the hypothesis that M_δ contains the constant functions and under the following regularity assumption.

Assumption 1 There exists a constant $\sigma > 0$ independent of δ such that, $\forall E \in \mathcal{T}_\delta$, the distance between any two vertices of E is larger than or equal to σh_E, where h_E is the diameter of E.

Indeed, since $v_\delta \mapsto \sum_{i\in\mathcal{I}} \left(\mathrm{K}_i \nabla v_{\delta i}, \nabla v_{\delta i}\right)_{F_i}$ is a norm on

$$W_\delta := \left\{v_\delta \in V_\delta \colon b\left(v_\delta, \psi_\delta\right) = 0 \quad \forall \psi_\delta \in M_\delta\right\}\,,$$

by (6), a_δ is coercive of W_δ. Moreover, by classical arguments (see [21, 40]), since VEM functions are piecewise polynomials on traces and Assumption 1 allows us to build a regular triangulation inside each polygon, there exists a constant $\beta > 0$ independent of δ such that the following *inf–sup* condition holds:

$$\inf_{\psi_\delta \in M_\delta} \sup_{v_\delta \in M_\delta} \frac{b\left(v_\delta, \psi_\delta\right)}{\left\|v_\delta\right\|_{V_\delta} \left\|\psi_\delta\right\|_{M_\delta}} \geq \beta\,.$$

An advantage of the present method, as opposite to the globally conforming case, is that $\lambda_{\delta m}$ is now a piecewise polynomial approximation of $\left[\!\left[\frac{\partial h_i}{\partial \hat{n}_m^i}\right]\!\right]_{\Gamma_m}$, whereas in the previous case the latter is obtained by evaluating the gradient of the numerical solution, and this may yield less accurate approximations.

3.4.1 Implementation

As previously sketched, the space $M_{\delta m}$, $m \in \mathcal{M}$, will contain piecewise polynomial (not necessarily continuous) functions, defined on the discretization of the trace Γ_m induced by $\mathcal{T}_{\delta i}$, with $\mathcal{I}_m = (i, j)$. Let N_h and N_λ be the total number of degrees of freedom of h_δ^0 and λ_δ, respectively. The vectors $\mathbf{h}$ and $\boldsymbol{\lambda}$ containing the degrees of freedom are the solution of the system

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{d} \end{pmatrix} ,$$

where $A \in \mathbb{R}^{N_h \times N_h}$ is defined, as in Sect. 3.3.1, as the block-diagonal matrix of the local stiffness matrices and $B \in \mathbb{R}^{N_\lambda \times N_h}$ collects the terms coming from b:

$$B_{lk} := b\left(\phi_k, \psi_l\right) \qquad \forall l \in \{1, \ldots, N_\lambda\},\ k \in \{1, \ldots, N_h\},$$

being ϕ_k the k-th basis function of V_δ and μ_l the l-th basis function of M_δ. Regarding the right-hand-side, we have

$$\begin{aligned} f_k &:= (f, \phi_k)_\delta + \left(h^N, \phi_k\right)_{\Gamma^N} - a_\delta\left(\mathcal{R}_\delta\left(h^D\right), \phi_l\right) && \forall k \in \{1, \ldots, N_h\}, \\ d_l &:= -b\left(\mathcal{R}_\delta\left(h^D\right), \psi_l\right) && \forall l \in \{1, \ldots, N_\lambda\}. \end{aligned}$$

4 Numerical Results

While referring the reader to [10, 11] for numerical results that validate the two proposed approaches, here we focus on a comparison between the two.

We consider two DFNs. The first example is a more simple DFN (Fig. 6a) presenting all possible configurations (tips, three fractures intersection point and traces spanning all the fracture) for which we have an analytical solution, reported in [11]. The second one (Fig. 6b) is a more complex DFN stochastically generated, containing 27 fractures and all common geometrical complexities. More details can be found in [10].

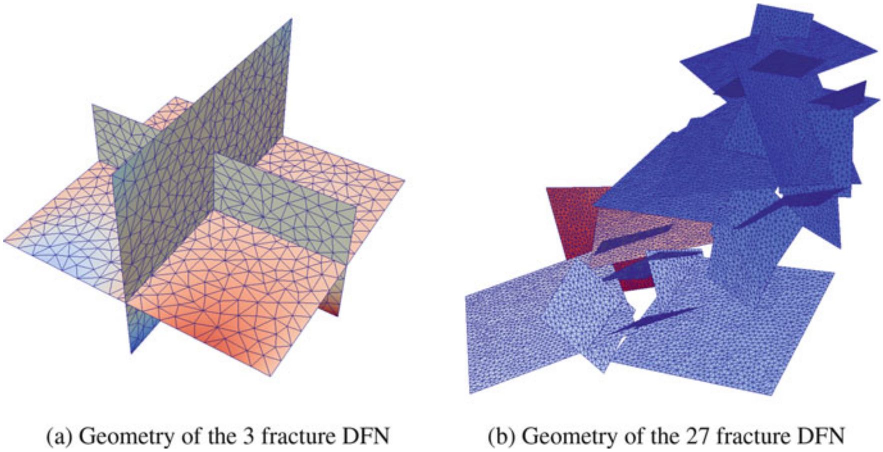

(a) Geometry of the 3 fracture DFN (b) Geometry of the 27 fracture DFN

Fig. 6 (**a**) Geometry of the 3 fracture DFN. (**b**) Geometry of the 27 fracture DFN

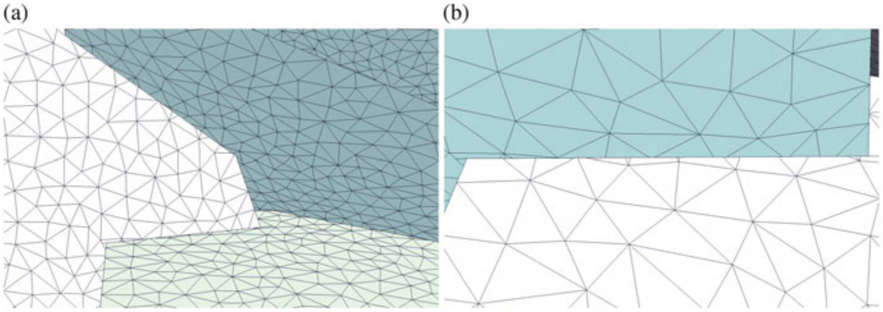


Fig. 7 (**a**) Non-conformity of the fracture meshes to traces. (**b**) Detail of non-conformity at fracture intersection.

Figure 7a displays a possible mesh configuration on the DFN in Fig. 6b. In Fig. 7b a detail of the previous mesh is reported. We highlight that, as said in Sect. 3.1, a triangular mesh is built independently on each fracture and then triangles are cut in along the traces, thus obtaining convex polygons. In the case of the globally conforming approach (Sect. 3.3), all the mesh points on the traces become degrees of freedom of the VEM spaces defined on each fracture (each physical point corresponds to two different degrees of freedom, one for each fracture). Instead, the VEM-Mortar approach (Sect. 3.4) requires that the points on the traces become degrees of freedom of the VEM space defined only on the fracture which they belong to.

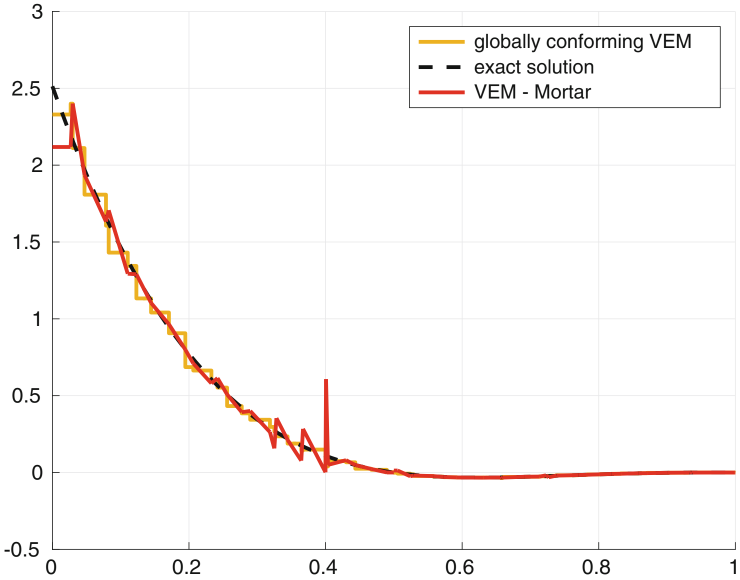


Fig. 8 Superposition of the fluxes computed by the two methods and the exact solution on a trace of the DFN in Fig. 6a

To compare the results obtained by the two methods, we focus on the fluxes computed on traces. Using the Mortar approach with piecewise linear Lagrange multipliers, we obtain an approximation of the flux at each trace from the solution λ_δ of the problem. In the globally conforming case, we obtain an approximation of the flux by post-processing the discrete solution, first projecting it on polynomials on each VEM element and then computing the jump of the co-normal derivative of the projection.

In Fig. 8 we compare the fluxes obtained by the conforming and the Mortar approaches to the exact solution for the trace with the tip (the vertical one) of the DFN in Fig. 6a. We can see a good agreement between the fluxes.

In Fig. 9 we show this comparison on three different traces of the DFN in Fig. 6b, from which we see that the two proposed approaches give comparable results.

In the Mortar approach the plot of the fluxes displays evident oscillations, due to the fact that convergence to the exact solution is proven in the $\mathrm{H}^{-\frac{1}{2}}(\Gamma_m)$-norm, $\forall m \in \mathcal{M}$. Nevertheless, the quality of the approximation is usually more accurate and reliable being the result of a direct computation and not of a post-processing process.

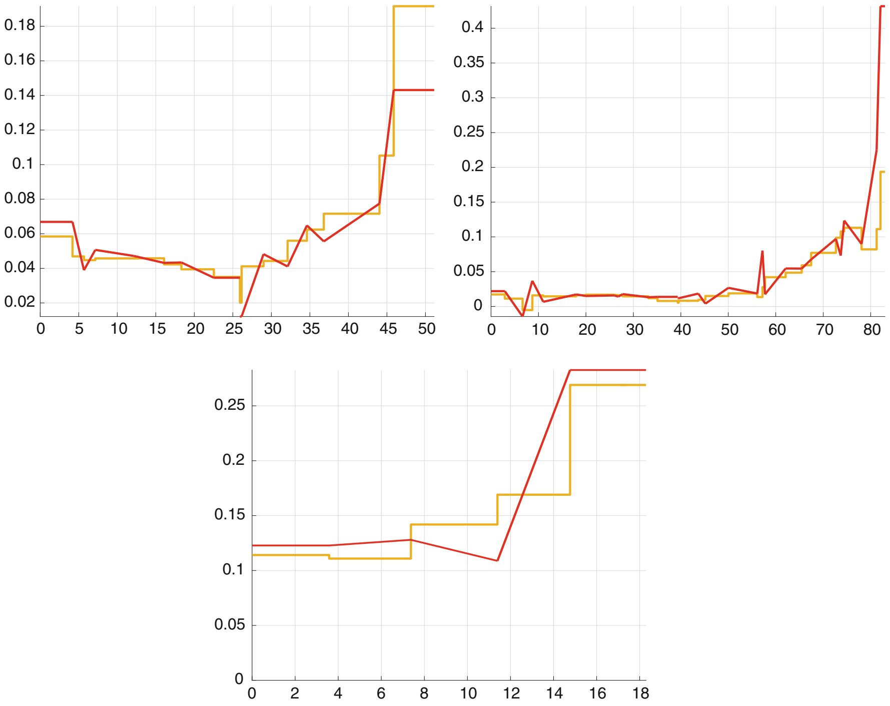


Fig. 9 Superposition of the fluxes computed by the two methods on three traces of the 27 fracture DFN in Fig. 6b

5 Conclusions

We have considered the application of the Virtual Element Methods to underground flow simulations. Among the several open and challenging issues stated by these applications there are the geometrical complexities that can pose very difficult problems in the mesh generation process. These difficulties are very strong constraints in the classical Finite Element approach to the simulations. The large flexibility introduced by the VEM in the mesh generation is a key issue that can largely help simulations.

We have discussed how it is possible to take advantage from this flexibility in practical applications. Numerical examples here presented, and more detailed numerical examples presented in [10, 11], clearly indicate the viability of the approach.

Acknowledgements This work has been supported by the Italian MIUR through PRIN research grant 2012HBLYE4_001 "Metodologie innovative nella modellistica differenziale numerica" and by INdAM-GNCS. Fernando Matias Benedetto was supported by the European Commission through the Erasmus Mundus Action 2-Strand1 ARCOIRIS programme, Politecnico di Torino. Stefano Berrone acknowledges the partial financial support received from the European Unions

Seventh Framework Programme (FP7/ 20142016) under Grant Agreement Number 607626 (Safeciti). Project title: “Simulation Platform for the Analysis of Crowd Turmoil in Urban Environments with Training and Predictive Capabilities”. This publication reflects the views only of the authors and the Commission cannot be held responsible for any use which may be made of the information here contained.

References

1. Adler, P.M.: Fractures and Fracture Networks. Kluwer Academic, Dordrecht (1999)
2. Ahmad, B., Alsaedi, A., Brezzi, F., Marini, L., Russo, A.: Equivalent projectors for virtual element methods. Comput. Math. Appl. **66**, 376–391 (2013)
3. Beirão da Veiga, L., Brezzi, F., Cangiani, A., Manzini, G., Marini, L., Russo, A.: Basic principles of virtual element methods. Math. Models Methods Appl. Sci. **23**(01), 199–214 (2013)
4. Beirão da Veiga, L., Brezzi, F., Marini, L.: Virtual elements for linear elasticity problems. SIAM J. Numer. Anal. **51**(2), 794–812 (2013)
5. Beirão Da Veiga, L., Brezzi, F., Marini, L., Russo, A.: The hitchhiker’s guide to the virtual element method. Math. Models Methods Appl. Sci **24**(8), 1541–1573 (2014)
6. Beirão da Veiga, L., Brezzi, F., Marini, L., Russo, A.: Virtual element methods for general second order elliptic problems on polygonal meshes. Available online at http://arxiv.org/abs/1412.2646 (2014)
7. Beirão da Veiga, L., Lovadina, C., Mora, D.: A virtual element method for elastic and inelastic problems on polytope meshes (2015). http://arxiv.org/abs/1503.02042. Available online
8. Belgacem, F.B.: The mortar finite element method with lagrange multipliers. Numer. Math. **84**(2), 173–197 (1999)
9. Benedetto, M.F., Berrone, S., Pieraccini, S., Scialò, S.: The virtual element method for discrete fracture network simulations. Comput. Methods Appl. Mech. Eng. **280**(0), 135–156 (2014). doi:http://dx.doi.org/10.1016/j.cma.2014.07.016
10. Benedetto, M.F., Berrone, S., Scialò, S.: A globally conforming method for solving flow in discrete fracture networks using the virtual element method. Finite Elem. Anal. Des. **109**, 23–36 (2016). doi:http://dx.doi.org/10.1016/j.finel.2015.10.003
11. Benedetto, M., Berrone, S., Borio, A., Pieraccini, S., Scialò, S.: A hybrid mortar virtual element method for discrete fracture network simulations. J. Comput. Phys. **306**, 148–166 (2016). doi:http://dx.doi.org/10.1016/j.jcp.2015.11.034
12. Bernardi, C., Maday, Y., Patera, A.T.: A new nonconforming approach to domain decomposition: the mortar element method. In: Nonlinear Partial Differential Equations and Their Applications. Collège de France Seminar, vol. XI (Paris, 1989–1991), Pitman Res. Notes Math. Ser., vol. 299, pp. 13–51. Longman Sci. Tech., Harlow (1994)
13. Bernardi, C., Maday, Y., Rapetti, F.: Basics and some applications of the mortar element method. GAMM-Mitteilungen **28**(2), 97–123 (2005)
14. Berrone, S., Pieraccini, S., Scialò, S.: A PDE-constrained optimization formulation for discrete fracture network flows. SIAM J. Sci. Comput. **35**(2), B487–B510 (2013). doi:http://dx.doi.org/10.1137/120865884
15. Berrone, S., Pieraccini, S., Scialò, S.: On simulations of discrete fracture network flows with an optimization-based extended finite element method. SIAM J. Sci. Comput. **35**(2), A908–A935 (2013). doi:http://dx.doi.org/10.1137/120882883
16. Berrone, S., Pieraccini, S., Scialò, S.: An optimization approach for large scale simulations of discrete fracture network flows. J. Comput. Phys. **256**, 838–853 (2014). doi:http://dx.doi.org/10.1016/j.jcp.2013.09.028

17. Berrone, S., Pieraccini, S., Scialò, S., Vicini, F.: A parallel solver for large scale DFN flow simulations. SIAM J. Sci. Comput. **37**(3), C285–C306 (2015). doi:http://dx.doi.org/10.1137/140984014
18. Berrone, S., Canuto, C., Pieraccini, S., Scialò, S.: Uncertainty quantification in discrete fracture network models: stochastic fracture transmissivity. Comput. Math. Appl. **70**(4), 603–623 (2015). doi:http://dx.doi.org/10.1016/j.camwa.2015.05.013
19. Berrone, S., Borio, A., Scialò, S.: A posteriori error estimate for a PDE-constrained optimization formulation for the flow in DFNs. SIAM J. Numer. Anal. **54**(1), 242–261 (2016)
20. Berrone, S., Pieraccini, S., Scialò, S.: Towards effective flow simulations in realistic discrete fracture networks. J. Comput. Phys. **310**, 181–201 (2016). doi:http://dx.doi.org/10.1016/j.jcp.2016.01.009
21. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. Rev. Française Automat. Informat. Recherche opérationelle. Anal. Numér. **8**(2), 129–151 (1974)
22. Brezzi, F., Falk, R.S., Marini, D.L.: Basic principles of mixed virtual element methods. ESAIM Math. Modell. Numer. Anal. **48**(04), 1227–1240 (2014)
23. Cammarata, G., Fidelibus, C., Cravero, M., Barla, G.: The hydro-mechanically coupled response of rock fractures. Rock Mech. Rock. Eng. **40**(1), 41–61 (2007). doi:http://dx.doi.org/10.1007/s00603-006-0081-z
24. Dershowitz, W.S., Fidelibus, C.: Derivation of equivalent pipe networks analogues for three-dimensional discrete fracture networks by the boundary element method. Water Resour. Res. **35**, 2685–2691 (1999). doi:http://dx.doi.org/10.1029/1999WR900118
25. de Dreuzy, J.R., Pichot, G., Poirriez, B., Erhel, J.: Synthetic benchmark for modeling flow in 3D fractured media. Comput. Geosci. **50**(0), 59–71 (2013)
26. Fidelibus, C., Cammarata, G., Cravero, M.: Hydraulic characterization of fractured rocks. In: Abbie, M., Bedford, J.S. (eds.) Rock Mechanics: New Research. Nova Science Publishers Inc., New York (2009)
27. Formaggia, L., Fumagalli, A., Scotti, A., Ruffo, P.: A reduced model for Darcy's problem in networks of fractures. ESAIM Math. Modell. Numer. Anal. **48**, 1089–1116 (2014). doi:http://dx.doi.org/0.1051/m2an/2013132
28. Formaggia, L., Antonietti, P., Panfili, P., Scotti, A., Turconi, L., Verani, M., Cominelli, A.: Optimal techniques to simulate flow in fractured reservoir. In: ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery (2014)
29. Fumagalli, A., Scotti, A.: A numerical method for two-phase flow in fractured porous media with non-matching grids. Adv. Water Resour. **62**, 454–464 (2013). doi:http://dx.doi.org/10.1016/j.advwatres.2013.04.001
30. Hyman, J., Gable, C., Painter, S., Makedonska, N.: Conforming Delaunay triangulation of stochastically generated three dimensional discrete fracture networks: a feature rejection algorithm for meshing strategy. SIAM J. Sci. Comput. **36**, A1871–A1894 (2014). doi:http://dx.doi.org/10.1137/130942541
31. Jaffré, J., Roberts, J.E.: Modeling flow in porous media with fractures; discrete fracture models with matrix-fracture exchange. Numer. Anal. Appl. **5**(2), 162–167 (2012)
32. Klawonn, A.: FETI domain decomposition methods for second order elliptic partial differential equations. GAMM-Mitteilungen **29**(2), 319–341 (2006). doi:10.1002/gamm.201490036
33. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Berlin (1999)
34. Nœtinger, B.: A quasi steady state method for solving transient Darcy flow in complex 3D fractured networks accounting for matrix to fracture flow. J. Comput. Phys. **283**, 205–223 (2015). doi:http://dx.doi.org/10.1016/j.jcp.2014.11.038
35. Nœtinger, B., Jarrige, N.: A quasi steady state method for solving transient Darcy flow in complex 3D fractured networks. J. Comput. Phys. **231**(1), 23–38 (2012). doi:http://dx.doi.org/10.1016/j.jcp.2011.08.015
36. Pichot, G., Erhel, J., de Dreuzy, J.: A mixed hybrid mortar method for solving flow in discrete fracture networks. Appl. Anal. **89**, 1629–643 (2010). doi:http://dx.doi.org/10.1080/00036811.2010.495333

37. Pichot, G., Erhel, J., de Dreuzy, J.: A generalized mixed hybrid mortar method for solving flow in stochastic discrete fracture networks. SIAM J. Sci. Comput. **34**, B86–B105 (2012). doi:http://dx.doi.org/10.1137/100804383
38. Pichot, G., Poirriez, B., Erhel, J., de Dreuzy, J.R.: A Mortar BDD method for solving flow in stochastic discrete fracture networks. In: Domain Decomposition Methods in Science and Engineering XXI. Lecture Notes in Computational Science and Engineering, pp. 99–112. Springer, Cham (2014)
39. Pieraccini, S., Scialò, S.: On a PDE-constrained optimization approach for flow simulations in fractured media. In: Ventura, G., Benvenuti, E. (eds.) Advances in Discretization Methods. SEMA SIMAI Springer Series vol. 12. Springer International Publishing, Cham (2016)
40. Raviart, P., Thomas, J.: Primal hybrid finite element methods for 2nd order elliptic equations. Math. Comput. **31**(138), 391–413 (1977)
41. Vohralík, M., Maryška, J., Severýn, O.: Mixed and nonconforming finite element methods on a system of polygons. Appl. Numer. Math. **51**, 176–193 (2007)
42. Wohlmuth, B.I.: A mortar finite element method using dual spaces for the lagrange multiplier. SIAM J. Numer. Anal. **38**(3), 989–1012 (2000)
43. Wohlmuth, B.I.: Discretization Techniques Based on Domain Decomposition. Springer, Berlin (2001)

Adaptive Discontinuous Galerkin Methods on Polytopic Meshes

Joe Collis and Paul Houston

Abstract In this article we consider the application of discontinuous Galerkin finite element methods, defined on agglomerated meshes consisting of general polytopic elements, to the numerical approximation of partial differential equation problems posed on complicated geometries. Here, we assume that the underlying computational domain may be accurately represented by a geometry-conforming fine mesh $\mathscr{T}_{\text{fine}}$; the resulting coarse mesh is then constructed based on employing standard graph partitioning algorithms. To improve the accuracy of the computed numerical approximation, we consider the development of goal-oriented adaptation techniques within an automatic mesh refinement strategy. In this setting, elements marked for refinement are subdivided by locally constructing finer agglomerates; should further resolution of the underlying fine mesh $\mathscr{T}_{\text{fine}}$ be required, then adaptive refinement of $\mathscr{T}_{\text{fine}}$ will also be undertaken. As an example of the application of these techniques, we consider the numerical approximation of the linear elasticity equations for a homogeneous isotropic material. In particular, the performance of the proposed adaptive refinement algorithm is studied for the computation of the (scaled) effective Young's modulus of a section of trabecular bone.

1 Introduction

Over the last couple of decades extensive work has been undertaken on the design and mathematical analysis of numerical methods for the approximation of partial differential equations (PDEs) based on exploiting general meshes consisting of polytopic elements, i.e., polygons/polyhedra in two-/three-dimensions, respectively. In particular, we mention the Polygonal Finite Element Method [34], the Extended Finite Element Method [21], the Mimetic Finite Difference Method [10, 12, 16], the Virtual Element Method [11], the Hybrid High Order Method [19, 20], the Composite Finite Element Method [1, 24–26], and the closely related Agglomerated Discontinuous Galerkin (DG) method [5–7]. The exploitation of general polytopic

---

J. Collis • P. Houston (✉)
School of Mathematical Sciences, University of Nottingham, University Park,
Nottingham NG7 2RD, UK
e-mail: Joe.Collis@nottingham.ac.uk; Paul.Houston@nottingham.ac.uk

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_9

elements offers great flexibility for mesh generation, and moreover allows for sequences of nested, successively coarser, meshes to be generated for use within multi-level solvers, such as multigrid and domain decomposition preconditioners, cf. [2, 4, 8, 22], for example. We point out that polytopic elements naturally arise when fictitious domain methods, unfitted methods or overlapping meshes are employed, cf. [13–15, 28, 30], for example.

The motivation here for employing polytopic elements is very much inspired by the work undertaken by Hackbusch and Sauter on Composite Finite Element methods in the articles [25, 26]; for the extension to DG methods, we refer to [1, 24]. Here, polytopic elements allow for the construction of a geometry-conforming mesh with a very small number of elements, irrespective of the complexity of the underlying domain; as an example in Fig. 1 we consider a section of trabecular bone which will be treated in Sect. 5.2 below. Indeed, by removing the need to have standard-shaped elements, i.e., triangles/quadrilaterals in two-dimensions and tetrahedra/hexahedra/prisms/pyramids in three-dimensions, highly complex geometries may be accurately meshed using a small number of (polytopic) elements. In the series of articles [1, 24–26], the underlying mesh is constructed based on adaptively refining an overlapping mesh and removing elements which do not lie inside the domain until a suitably accurate representation of the given geometry is computed. By exploiting the underlying tree structure generated by this adaptive refinement procedure, general polytopic elements are formed by agglomerating elements which share the same parent; for a review of this approach in the DG setting, we refer to [3]. The generation of such coarse meshes is advantageous from a computational point of view, in the sense that coarse approximations may be computed very efficiently;

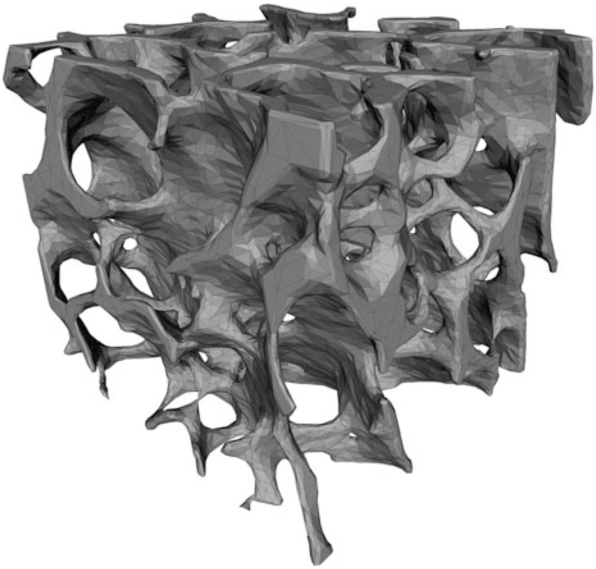


Fig. 1 Section of trabecular bone, cf. [31]

however, the accuracy of the resulting numerical solution may be insufficient and subsequent mesh refinement may be desirable. In the composite finite element framework adaptive mesh refinement may be undertaken in a simple manner by constructing finite element partitions of the domain consisting of agglomerated elements which belong to different levels of the underlying hierarchical tree data structure; indeed, this approach has been exploited within the articles [23, 24].

In this article we consider an alternative approach whereby the underlying geometry-conforming mesh may be constructed more generally using, for example, standard mesh generation software, rather than the refined-overlapping mesh approach considered in [23, 24]. The underlying coarse composite mesh is then constructed using standard graph partitioning algorithms; for this purpose, here we employ METIS [29]. On the basis of this coarse agglomerated mesh partition we consider the application of goal-oriented dual-weighted-residual (DWR) a posteriori error estimation for DG finite element methods; to fix ideas, we focus on the discretization of the linear elasticity equations for a homogeneous isotropic material. However, we stress that this approach is completely general and can be applied to general classes of PDE problems. Once elements have been marked for refinement on the basis of the DWR error indicators, the agglomerated elements can be subdivided by simply employing a local graph partitioning algorithm involving the set of (standard) elements which form the marked polytope. In this way adaptive refinement can easily and efficiently be implemented without the need for complicated mesh refinement tree data structures. However, once a refined element is of the granularity of an element in the background fine mesh, then standard refinement of the fine mesh may then need to be undertaken. To illustrate the performance of the proposed agglomeration-based refinement algorithm, we present a series of numerical examples; in particular, we consider the aforementioned three-dimensional section of trabecular bone depicted in Fig. 1.

The outline of this article is as follows. In Sect. 2 we introduce, as a prototype PDE problem, the linear elasticity equations for a homogeneous isotropic material. Section 3 is then devoted to the formulation of the corresponding DG discretization, based on employing the symmetric version of the interior penalty method, cf. [27, 36]. In Sect. 4 we briefly introduce a goal-oriented a posteriori error estimator, followed by an outline of the design of an appropriate agglomeration-based adaptive mesh refinement algorithm. The practical performance of the proposed adaptive refinement strategy is studied in Sect. 5. Finally, in Sect. 6 we summarize the work presented in this article and draw some conclusions.

2 Model Problem

Given that Ω is a bounded, connected Lipschitz domain in $\mathbb{R}^d$, $d > 1$, with boundary $\partial\Omega$, consider the following linear elasticity equation: find $\mathbf{u}$ such that

$$-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}) = \mathbf{f} \ \text{ in } \Omega, \tag{1}$$

where $\mathbf{u} = (u_1, \ldots, u_d)^\top$ is the displacement and $\boldsymbol{\sigma}$ is the stress tensor for a homogeneous isotropic material, i.e.,

$$\boldsymbol{\sigma}(\mathbf{u}) = 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) + \lambda\nabla\cdot\mathbf{u}I,$$

I is the $d \times d$ identity matrix, $\boldsymbol{\varepsilon}(\mathbf{u}) = 1/2(\nabla\mathbf{u} + \nabla\mathbf{u}^\top)$, and μ and λ are the Lamé coefficients, which satisfy the relation

$$0 < \min\{\mu, \mu + \lambda\}.$$

We divide $\partial\Omega$ into the disjoint subsets $\partial\Omega_{\mathrm{D}}$, $\partial\Omega_{\mathrm{ND}}$, and $\partial\Omega_{\mathrm{N}}$ whose union is $\partial\Omega$, with $\partial\Omega_{\mathrm{D}}$ or $\partial\Omega_{\mathrm{ND}}$ nonempty and relatively open in $\partial\Omega$. Following [35], we supplement (1) with the following boundary conditions

$$\begin{aligned} \mathbf{u} &= \mathbf{g}_{\mathrm{D}} && \text{on } \partial\Omega_{\mathrm{D}}, \\ \mathbf{u}\cdot\mathbf{n} &= g_{\mathrm{ND}} && \text{on } \partial\Omega_{\mathrm{ND}}, \\ \boldsymbol{\sigma}(\mathbf{u})\mathbf{n}\cdot\mathbf{t} &= 0 && \text{on } \partial\Omega_{\mathrm{ND}}, \\ \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} &= \mathbf{g}_{\mathrm{N}} && \text{on } \partial\Omega_{\mathrm{N}}, \end{aligned} \tag{2}$$

where $\mathbf{n}$ and $\mathbf{t}$ denote the unit outward normal vector and unit tangential vector(s) on the boundary $\partial\Omega$.

3 Interior Penalty Discontinuous Galerkin Method

In this section we introduce the DG discretization of the model problem (1), (2) based on employing the (symmetric) version of the interior penalty method, cf. [27, 36].

To this end, let $\mathscr{T}$ be a subdivision of the computational domain Ω into disjoint open polytopic elements κ such that $\bar{\Omega} = \cup_{\kappa\in\mathscr{T}}\bar{\kappa}$. For the purposes of this article the polytopic mesh $\mathscr{T}$ will be constructed based on an agglomeration of a geometry-conforming fine mesh $\mathscr{T}_{\text{fine}}$ consisting of standard element types, i.e., triangles/quadrilaterals in two-dimensions, and so on, cf. Sect. 1 above. Further details concerning the construction of $\mathscr{T}$ will be given below in Sect. 4. On the basis of the (polygonal/polyhedral) mesh $\mathscr{T}$, given the polynomial degree $p \geq 1$, we define the corresponding DG finite element space $\mathbf{V}_p(\mathscr{T})$ by

$$\mathbf{V}_p(\mathscr{T}) = \{\mathbf{u} \in [L_2(\Omega)]^d : \mathbf{u}|_\kappa \in [\mathscr{P}_p(\kappa)]^d,\ \ \kappa \in \mathscr{T}\},$$

where $\mathscr{P}_p(\kappa)$ denotes the space of polynomials of total degree p over κ. As in [17], the local elemental polynomial spaces employed within the definition of $\mathbf{V}_p(\mathscr{T})$ are constructed in the physical space, without the need to map from a given reference or canonical frame.

Following [17], we define the *interfaces* of the mesh $\mathscr{T}$ to be the set of $(d-1)$-dimensional facets of the elements $\kappa \in \mathscr{T}$. To allow for the presence of hanging nodes/edges, the interfaces of $\mathscr{T}$ are defined to be the intersection of the $(d-1)$-dimensional facets of neighbouring elements. In two-dimensions, i.e., when $d = 2$, the interfaces of a given element $\kappa \in \mathscr{T}$ will always consist of $(d-1)$-dimensional simplices (line segments). In general, for $d = 3$, this will not be the case; in this setting, we assume that each interface of an element $\kappa \in \mathscr{T}$ may be subdivided by a set of co-planar triangles. Thereby, we use the terminology 'face' to refer to a $(d-1)$-dimensional simplex, which forms part of the boundary (interface) of an element $\kappa \in \mathscr{T}$.

As in [17, 18], we assume that a sub-triangulation into faces of each mesh interface is given if $d = 3$, and denote by $\mathscr{F}$ the union of all open mesh interfaces if $d = 2$ and the union of all open triangles belonging to the sub-triangulation of all mesh interfaces if $d = 3$. We note that this assumption is trivially satisfied in our setting when $\mathscr{T}$ is formed by the agglomeration of a fine mesh $\mathscr{T}_{\text{fine}}$ consisting of simplices. We write $\mathscr{F} = \mathscr{F}^{\mathscr{I}} \cup \mathscr{F}^{\mathscr{B}}$, where $\mathscr{F}^{\mathscr{I}}$ denotes the set of all open $(d-1)$-dimensional element faces $F \in \mathscr{F}$ that are contained in Ω, and $\mathscr{F}^{\mathscr{B}}$ is the union of element boundary faces. Furthermore, we write $\mathscr{F}^{\mathscr{B}} = \mathscr{F}^{\mathscr{B}}_{\mathrm{D}} \cup \mathscr{F}^{\mathscr{B}}_{\mathrm{ND}} \cup \mathscr{F}^{\mathscr{B}}_{\mathrm{N}}$ where $\mathscr{F}^{\mathscr{B}}_{\mathrm{D}}$, $\mathscr{F}^{\mathscr{B}}_{\mathrm{ND}}$, and $\mathscr{F}^{\mathscr{B}}_{\mathrm{N}}$ denote the set of boundary faces whose union form $\partial\Omega_{\mathrm{D}}$, $\partial\Omega_{\mathrm{ND}}$, and $\partial\Omega_{\mathrm{N}}$, respectively. Here, the boundary $\partial\kappa$ of an element κ and the sets $\partial\kappa\backslash\partial\Omega$, $\partial\kappa \cap \partial\Omega_{\mathrm{D}}$, $\partial\kappa \cap \partial\Omega_{\mathrm{ND}}$, and $\partial\kappa \cap \partial\Omega_{\mathrm{N}}$ will be identified in a natural way with the corresponding subsets of $\mathscr{F}$. Implicit in these definitions is the assumption that $\mathscr{T}$ respects the decomposition of $\partial\Omega$ in the sense that each $F \in \mathscr{F}^{\mathscr{B}}$ belongs to the interior of exactly one of $\partial\Omega_{\mathrm{D}}$, $\partial\Omega_{\mathrm{ND}}$, or $\partial\Omega_{\mathrm{N}}$.

Next, we define average and jump operators. To this end, let κ^+ and κ^- be two adjacent elements of $\mathscr{T}$, and $\mathbf{x}$ be an arbitrary point on the interior face $F \subset \partial\kappa^+ \cap \partial\kappa^-$, $F \in \mathscr{F}^{\mathscr{I}}$. Given vector- and matrix-valued functions $\mathbf{v}$ and $\boldsymbol{\tau}$, respectively, that are smooth inside each element $\kappa^\pm$, by $(\mathbf{v}^\pm, \boldsymbol{\tau}^\pm)$ we denote the traces of $(\mathbf{v}, \boldsymbol{\tau})$ on F taken from within the interior of $\kappa^\pm$, respectively. Then, we introduce the averages at $\mathbf{x} \in F$:

$$\{\!\{\mathbf{v}\}\!\} = (\mathbf{v}^+ + \mathbf{v}^-)/2, \quad \{\!\{\boldsymbol{\tau}\}\!\} = (\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-)/2.$$

Similarly, the jumps of $\mathbf{v}$ at $\mathbf{x} \in F$ are given by

$$\underline{[\![\mathbf{v}]\!]} = \mathbf{v}^+ \otimes \mathbf{n}_{\kappa^+} + \mathbf{v}^- \otimes \mathbf{n}_{\kappa^-}, \quad [\![\mathbf{v}]\!] = \mathbf{v}^+ \cdot \mathbf{n}_{\kappa^+} + \mathbf{v}^- \cdot \mathbf{n}_{\kappa^-}.$$

On a boundary face $F \in \mathscr{F}^{\mathscr{B}}$, we set $\{\!\{\mathbf{v}\}\!\} = \mathbf{v}$, $\{\!\{\boldsymbol{\tau}\}\!\} = \boldsymbol{\tau}$, $\underline{[\![\mathbf{v}]\!]} = \mathbf{v} \otimes \mathbf{n}$, and $[\![\mathbf{v}]\!] = \mathbf{v} \cdot \mathbf{n}$, where $\mathbf{n}$ denotes the unit outward normal vector on the boundary $\partial\Omega$.

With this notation, the symmetric version of the interior penalty DG method is given by: find $\mathbf{u}_h \in \mathbf{V}_p(\mathscr{T})$ such that

$$B(\mathbf{u}_h, \mathbf{v}_h) = \ell(\mathbf{v}_h) \tag{3}$$

for all $\mathbf{v}_h \in \mathbf{V}_p(\mathscr{T})$. Here, the bilinear form $B : \mathbf{V}_p(\mathscr{T}) \times \mathbf{V}_p(\mathscr{T}) \to \mathbb{R}$ is given by

$$\begin{aligned}
B(\mathbf{w}, \mathbf{v}) := &\sum_{\kappa \in \mathscr{T}} \int_\kappa \boldsymbol{\sigma}(\mathbf{w}) : \boldsymbol{\varepsilon}(\mathbf{v}) \,\mathrm{d}\mathbf{x} \\
&- \sum_{F \in \mathscr{F}^{\mathscr{I}} \cup \mathscr{F}^{\mathscr{B}}_{\mathrm{D}}} \int_F \Big(\{\!\{\boldsymbol{\sigma}_h(\mathbf{w})\}\!\} : \underline{[\![\mathbf{v}]\!]} + \{\!\{\boldsymbol{\sigma}_h(\mathbf{v})\}\!\} : \underline{[\![\mathbf{w}]\!]} - \alpha\, \mu\, \underline{[\![\mathbf{w}]\!]} : \underline{[\![\mathbf{v}]\!]} \Big) \,\mathrm{d}s \\
&+ \sum_{F \in \mathscr{F}^{\mathscr{I}} \cup \mathscr{F}^{\mathscr{B}}_{\mathrm{D}} \cup \mathscr{F}^{\mathscr{B}}_{\mathrm{ND}}} \int_F \alpha\, \lambda\, [\![\mathbf{w}]\!][\![\mathbf{v}]\!] \,\mathrm{d}s \\
&- \sum_{F \in \mathscr{F}^{\mathscr{B}}_{\mathrm{ND}}} \int_F \left(((\boldsymbol{\sigma}_h(\mathbf{w})\mathbf{n}) \cdot \mathbf{n})\, (\mathbf{v} \cdot \mathbf{n}) + ((\boldsymbol{\sigma}_h(\mathbf{v})\mathbf{n}) \cdot \mathbf{n})\, (\mathbf{w} \cdot \mathbf{n}) \right) \,\mathrm{d}s,
\end{aligned}$$

and the linear functional $\ell : \mathbf{V}_p(\mathscr{T}) \to \mathbb{R}$ is defined by

$$\begin{aligned}
\ell(\mathbf{v}) = &\int_\Omega \mathbf{f} \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} - \int_{\partial\Omega_{\mathrm{D}}} \boldsymbol{\sigma}_h(\mathbf{v}) : \mathbf{g}_{\mathrm{D}} \otimes \mathbf{n} \,\mathrm{d}s + \int_{\partial\Omega_{\mathrm{D}}} \alpha\, \mu\, \mathbf{g}_{\mathrm{D}} \cdot \mathbf{v} \,\mathrm{d}s \\
&+ \int_{\partial\Omega_{\mathrm{D}}} \alpha\, \lambda\, (\mathbf{g}_{\mathrm{D}} \cdot \mathbf{n})(\mathbf{v} \cdot \mathbf{n}) \,\mathrm{d}s - \int_{\partial\Omega_{\mathrm{ND}}} g_{\mathrm{ND}}((\boldsymbol{\sigma}_h(\mathbf{v})\mathbf{n}) \cdot \mathbf{n} - \alpha\, \lambda\, \mathbf{v} \cdot \mathbf{n}) \,\mathrm{d}s \\
&+ \int_{\partial\Omega_{\mathrm{N}}} \mathbf{g}_{\mathrm{N}} \cdot \mathbf{v} \,\mathrm{d}s,
\end{aligned}$$

where $\boldsymbol{\sigma}_h$ is the stress tensor defined elementwise.

The non-negative function α is referred to as the *discontinuity-penalization parameter*; the precise definition will be given below based on the work undertaken in our recent article [17]. To this end, following [17] we first introduce the submesh $\tilde{\mathscr{T}}$ of elements from $\mathscr{T}$.

Definition 1 Let $\tilde{\mathscr{T}}$ denote the subset of elements κ, $\kappa \in \mathscr{T}$, such that each $\kappa \in \tilde{\mathscr{T}}$ can be covered by at most $n_{\mathscr{T}}$ shape-regular simplices $\mathscr{K}_i$, $i = 1, \ldots, n_{\mathscr{T}}$, such that

$$\mathrm{dist}(\kappa, \partial\mathscr{K}_i) > C_{as}\, \mathrm{diam}(\mathscr{K}_i)/p_\kappa^2, \quad \text{and} \quad |\mathscr{K}_i| \geq c_{as}|\kappa|$$

for all $i = 1, \ldots, n_{\mathscr{T}}$, for some $n_{\mathscr{T}} \in \mathbb{N}$ and $C_{as}, c_{as} > 0$, independent of κ and $\mathscr{T}$.

With this definition, we recall the following inverse inequality from [17]; we stress that this result is sharp with respect to both the polynomial order p, and moreover takes into account $(d-k)$-dimensional element facet degeneration, where $k = 1, 2, \ldots, d-1$.

Lemma 1 *Let $\kappa \in \mathscr{T}$, $F \subset \partial\kappa$ denote one of its faces, and $\tilde{\mathscr{T}}$ be defined as in Definition 1. Then, for each $v \in \mathscr{P}_p(\kappa)$, we have the inverse estimate*

$$\|v\|^2_{L_2(F)} \leq C_{\mathrm{INV}}(p,\kappa,F)\frac{p^2|F|}{|\kappa|}\|v\|^2_{L_2(\kappa)}, \tag{4}$$

with

$$C_{\mathrm{INV}}(p,\kappa,F) := C_{\mathrm{inv}}\begin{cases}\min\left\{\dfrac{|\kappa|}{\sup_{\kappa_\flat^F\subset\kappa}|\kappa_\flat^F|}, p^{2d}\right\}, & \text{if } \kappa \in \tilde{\mathscr{T}},\\ \dfrac{|\kappa|}{\sup_{\kappa_\flat^F\subset\kappa}|\kappa_\flat^F|}, & \text{if } \kappa \in \mathscr{T}\backslash\tilde{\mathscr{T}},\end{cases}$$

and $\kappa_\flat^F$ denotes a d-dimensional simplex contained in κ which shares the face F with $\kappa \in \mathscr{T}$. Furthermore, C_{inv} is a positive constant, which if $\kappa \in \tilde{\mathscr{T}}$ depends on the shape regularity of the covering of κ given in Definition 1, but is always independent of $|\kappa|/\sup_{\kappa_\flat^F\subset\kappa}|\kappa_\flat^F|$ (and, therefore, of $|F|$), p, and v.

Based on the inverse inequality stated in Lemma 1, together with the analysis presented in [17, 27] the discontinuity-penalization parameter α may be defined as follows.

Definition 2 Let $\alpha : \mathscr{F} \rightarrow \mathbb{R}_+$ be defined facewise by

$$\alpha(\mathbf{x}) = C_\alpha\begin{cases}\max\limits_{\kappa\in\{\kappa^+,\kappa^-\}}\left\{C_{\mathrm{INV}}(p_\kappa,\kappa,F)\dfrac{p^2|F|}{|\kappa|}\right\}, & \mathbf{x}\in F\in\mathscr{F}^{\mathscr{I}},\ F\subset\partial\kappa^+\cap\partial\kappa^-,\\ C_{\mathrm{INV}}(p_\kappa,\kappa,F)\dfrac{p^2|F|}{|\kappa|}, & \mathbf{x}\in F\in\mathscr{F}^{\mathscr{B}},\ F\subset\partial\kappa\cap\partial\Omega,\end{cases} \tag{5}$$

with $C_\alpha > C_\alpha^{\min}$, where $C_\alpha^{\min}$ is a sufficiently large lower bound.

4 Error Estimation and Adaptive Mesh Refinement

As noted above, for the purposes of this article we assume that the computational mesh $\mathscr{T}$, consisting of general polytopic elements, is constructed based on agglomerating an underlying fine mesh $\mathscr{T}_{\mathrm{fine}}$. Here, $\mathscr{T}_{\mathrm{fine}}$ may be constructed using a standard mesh generator; for example, we employ Triangle [32] and Tetgen [33] in two- and three-dimensions, respectively. By employing standard mesh generation software fine-scale geometry-conforming meshes may be generated; in the case when the computational domain Ω is highly complex, then $\mathscr{T}_{\mathrm{fine}}$ may consist of a very large number of elements, cf. [3], for example. With this mind, we employ METIS, cf. [29], for example, to construct $\mathscr{T}$ with a user-defined number of partitions based on agglomerating elements contained within $\mathscr{T}_{\mathrm{fine}}$. In general the construction of coarse computational meshes will not be sufficient to meet the accuracy demanded by a user. With this mind, we consider the design of a mesh refinement algorithm which automatically refines the agglomerates which form $\mathscr{T}$ based on repartitioning elements which possess a large error contribution. To illustrate this approach, we focus on employing a dual-weighted-residual (DWR) error estimator, cf. [9], for example, together with the references cited therein.

To this end, given a (linear, for simplicity) target functional J, we recall the following a posteriori error estimation formula

$$J(\mathbf{u}) - J(\mathbf{u}_h) = \ell(\mathbf{z} - \mathbf{z}_h) - B(\mathbf{u}_h, \mathbf{z} - \mathbf{z}_h) \equiv \sum_{\kappa \in \mathscr{T}} \eta_\kappa$$

for all $\mathbf{z}_h \in \mathbf{V}_p(\mathscr{T})$, where $\eta_\kappa = \eta_\kappa(\mathbf{u}_h, \mathbf{z} - \mathbf{z}_h)$ denotes the local elementwise error indicators on κ, $\kappa \in \mathscr{T}$. Furthermore, $\mathbf{z}$ denotes the solution of the corresponding dual/adjoint problem: find $\mathbf{z}$ such that

$$B(\mathbf{v}, \mathbf{z}) = J(\mathbf{v}) \quad \forall \mathbf{v}. \tag{6}$$

In practice, the error indicators η_κ, $\kappa \in \mathscr{T}$, are computed based on approximating the dual solution $\mathbf{z}$ by $\hat{\mathbf{z}}$ on the polytopic finite element mesh $\mathscr{T}$, using polynomials one degree higher than those employed for the computation of $\mathbf{u}_h$; thereby, $\hat{\mathbf{z}} \in \mathbf{V}_{p+1}(\mathscr{T})$. Writing $\hat{\eta}_\kappa = \eta_\kappa(\mathbf{u}_h, \hat{\mathbf{z}} - \mathbf{z}_h)$ gives rise to the approximate error representation formula

$$J(\mathbf{u}) - J(\mathbf{u}_h) \approx \sum_{\kappa \in \mathscr{T}} \hat{\eta}_\kappa.$$

On the basis of the size of the modulus of the (approximate) local error indicators, i.e., $|\hat{\eta}_\kappa|$, the elements in the mesh $\mathscr{T}$ are marked for refinement using the fixed fraction strategy with refinement parameter `REF`. Once an element $\kappa \in \mathscr{T}$ has been marked for refinement, then assuming that κ is formed from the union of a set of elements belonging to $\mathscr{T}_{\text{fine}} \equiv \mathscr{T}_{\text{fine}}^{(l)}$, $l = 0$, i.e., $\kappa = \cup_{\kappa' \in \mathscr{S}_\kappa^{(l)}} \kappa'$, where $\mathscr{S}_\kappa^{(l)} \subset \mathscr{T}$ denotes the set of fine elements which form κ, then METIS is applied to the corresponding graph representation of $\mathscr{S}_\kappa^{(l)}$ to yield a local partition of κ consisting of m_κ agglomerated elements; here, we set $m_\kappa = 2^d$. In the case when $m_\kappa > \operatorname{card}(\mathscr{S}_\kappa^{(l)})$ for any element $\kappa \in \mathscr{T}$ which has been marked for refinement, then the elements $\kappa' \in \mathscr{S}_\kappa^{(l)}$, $\kappa \in \mathscr{T}$, are first isotropically refined using standard adaptive mesh refinement algorithms to yield a new fine mesh $\mathscr{T}_{\text{fine}}^{(l+1)}$. Once $\mathscr{T}_{\text{fine}}^{(l+1)}$ has been constructed then new local partitions $\mathscr{S}_\kappa^{(l+1)}$ of each element $\kappa \in \mathscr{T}$ may be computed and, for those marked for refinement, subsequently subdivided using graph partitioning techniques. We stress that, assuming $\mathscr{T}_{\text{fine}}$ does not require adaptive mesh refinement to be undertaken, then the refinement of $\mathscr{T}$ can be done in a very straight-forward manner using only graph partitioning algorithms,

Algorithm 4.1 Construction and refinement of the agglomerated mesh $\mathscr{T}$

```
1:  Construct a geometry conforming fine mesh $\mathscr{T}_{\text{fine}} \equiv \mathscr{T}_{\text{fine}}^{(l)}$, $l = 0$, consisting of standard element shapes.
2:  Compute initial agglomerated mesh $\mathscr{T} \equiv \mathscr{T}^{(k)}$, $k = 0$, consisting of a user-defined number of elements.
3:  Solve (3) for $\mathbf{u}_h \in \mathbf{V}_p(\mathscr{T}^{(k)})$.
4:  Compute the numerical approximation $\hat{\mathbf{z}} \in \mathbf{V}_{p+1}(\mathscr{T}^{(k)})$ to the dual/adjoint problem (6).
5:  Evaluate the (approximate) element error indicators $\hat{\eta}_\kappa = \eta_\kappa(\mathbf{u}_h, \hat{\mathbf{z}} - \mathbf{z}_h)$, $\mathbf{z}_h \in \mathbf{V}_p(\mathscr{T}^{(k)})$, for each $\kappa \in \mathscr{T}^{(k)}$.
6:  if $\sum_{\kappa \in \mathscr{T}^{(k)}} \hat{\eta}_\kappa <$ TOL, where TOL is a user-defined tolerance then
7:      STOP
8:  end if
9:  Set refine_fine_mesh = False
10: Construct the refinement set $\mathscr{R}$ based on employing the fixed fraction refinement strategy.
11: for all $\kappa \in \mathscr{R}$ do
12:     if $m_\kappa > \text{card}(\mathscr{S}_\kappa^{(l)})$, where $m_\kappa = 2^d$ then
13:         for all $\kappa' \in \mathscr{S}_\kappa^{(l)}$ do
14:             Refine $\kappa'$ and set refine_fine_mesh = True
15:         end for
16:     end if
17: end for
18: if refine_fine_mesh then
19:     Set $l = l + 1$ and construct new fine mesh $\mathscr{T}_{\text{fine}}^{(l)}$.
20:     Recompute agglomeration partition for $\mathscr{T}^{(k)}$ relative to the new fine mesh $\mathscr{T}_{\text{fine}}^{(l)}$.
21: end if
22: for all $\kappa \in \mathscr{R}$ do
23:     Refine $\kappa$ based on computing a subpartition consisting of $m_\kappa$ elements.
24: end for
25: Set $k = k + 1$ and construct newly refined agglomerated mesh $\mathscr{T}^{(k)}$.
26: Goto 3.
```

without the need to implement complicated tree data structures, which are typically employed within standard refinement procedures. A summary of the proposed adaptive algorithm is presented in Algorithm 4.1.

5 Numerical Examples

In this section we present a series of numerical experiments to highlight the practical performance of the agglomeration-based adaptive mesh refinement algorithm outlined in Sect. 4. To this end, following the work presented in the recent article [35], we consider the evaluation of the (scaled) effective Young's modulus of a given structure; the key application presented in Sect. 5.2 is concerned with the modelling of the section of trabecular bone depicted in Fig. 1.

Throughout this section, we assume that the underlying geometry Ω is contained within a d-dimensional cuboid Ω_{cube}, where $\Omega_{\text{cube}} = \Pi_{i=1}^{d}(x_i^{\min}, x_i^{\max})$. Writing $\partial\Omega_{\text{cube}}$ to denote the boundary of Ω_{cube}, i.e., the planar sides of the cuboid, we define $\partial\Omega_{\mathrm{ND}} = \partial\Omega \cap \partial\Omega_{\mathrm{cube}}$ and $\partial\Omega_{\mathrm{N}} = \partial\Omega \backslash \partial\Omega_{\mathrm{ND}}$; thereby, $\partial\Omega_{\mathrm{D}} = \emptyset$. With this notation, we set $g_{\mathrm{ND}} = \bar{u} = 0.01 L_d$, $L_d = x_d^{\max} - x_d^{\min}$, on the top section of $\partial\Omega_{\mathrm{ND}}$, i.e., where $x_d = x_d^{\max}$, and $g_{\mathrm{ND}} = 0$ on all other portions of $\partial\Omega_{\mathrm{ND}}$. Furthermore, we set $\mathbf{g}_{\mathrm{N}} = \mathbf{0}$ on $\partial\Omega_{\mathrm{N}}$. Finally, writing E to denote the Young's modulus and ν the Poisson ratio, we define our functional of interest by

$$J(\mathbf{u}) = \frac{1}{E} \frac{L_d}{\bar{u} |\Omega_{\mathrm{cube}}|} \int_{\Omega} \sigma_{dd}(\mathbf{u}) \, \mathrm{d}\mathbf{x},$$

cf. [35].

5.1 Example 1: Two-Dimensional Cross Geometry

In this section we consider a simple two-dimensional cross-geometry enclosed within the unit square $(0,1)^2$; cf. Fig. 2. Here, we set $E = 10^6$ and the Poisson ratio $\nu = 0.3$; thereby, on the basis of a fine mesh calculation we compute an approximation to the (scaled) effective Young's modulus $J(\mathbf{u}) \approx 0.228481378$.

We generate an initial fine mesh $\mathscr{T}_{\mathrm{fine}}$ consisting of 111,457 triangular elements; this is then agglomerated using METIS to generate a coarse polytopic mesh $\mathscr{T}$ comprising 32 elements only, cf. Fig. 2. In Tables 1 and 2 we show the convergence

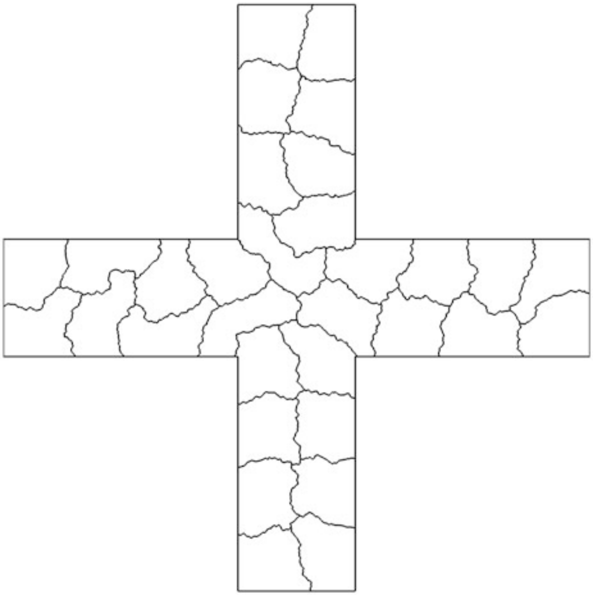


Fig. 2 Example 1. Initial agglomerated mesh consisting of 32 elements

Table 1 Example 1. Adaptive algorithm for $p = 1$

No of Eles	No of Dofs	$J(\mathbf{u}) - J(\mathbf{u}_h)$	$\sum_{\kappa \in \mathscr{T}} \hat{\eta}_\kappa$	θ
32	192	−3.794E−02	−3.142E−02	0.83
50	300	−2.424E−02	−1.973E−02	0.81
80	480	−1.561E−02	−1.307E−02	0.84
128	768	−1.075E−02	−9.145E−03	0.85
203	1218	−7.681E−03	−6.793E−03	0.88
323	1938	−5.535E−03	−5.100E−03	0.92
514	3084	−3.885E−03	−3.670E−03	0.94
820	4920	−2.661E−03	−2.563E−03	0.96
1308	7848	−1.759E−03	−1.707E−03	0.97
2084	12,504	−1.143E−03	−1.112E−03	0.97
3316	19,896	−7.352E−04	−7.182E−04	0.98
5274	31,644	−4.534E−04	−4.443E−04	0.98
8366	50,196	−2.717E−04	−2.680E−04	0.99
13,214	79,284	−1.565E−04	−1.551E−04	1.00
20,798	124,788	−8.965E−05	−8.946E−05	1.00

Table 2 Example 1. Adaptive algorithm for $p = 2$

No of Eles	No of Dofs	$J(\mathbf{u}) - J(\mathbf{u}_h)$	$\sum_{\kappa \in \mathscr{T}} \hat{\eta}_\kappa$	θ
32	384	−9.499E−03	−4.471E−03	0.47
50	600	−5.955E−03	−3.318E−03	0.56
80	960	−3.458E−03	−2.111E−03	0.61
128	1536	−2.149E−03	−1.473E−03	0.69
203	2436	−1.140E−03	−8.385E−04	0.74
323	3876	−5.537E−04	−4.417E−04	0.80
514	6168	−2.645E−04	−2.221E−04	0.84
817	9804	−1.114E−04	−9.954E−05	0.89
1301	15,612	−4.839E−05	−4.337E−05	0.90
2068	24,816	−2.051E−05	−1.839E−05	0.90
3277	39,324	−8.643E−06	−7.999E−06	0.93
5216	62,592	−3.491E−06	−3.491E−06	1.00

history of the proposed agglomeration-based adaptive strategy using a polynomial order of degree $p = 1$ and $p = 2$, respectively, with `REF` $= 20\,\%$. In each case, we show the number of elements in polytopic mesh $\mathscr{T}$, the number of degrees of freedom in underlying finite element space $\mathbf{V}_p(\mathscr{T})$, the true error in the (scaled) effective Young's modulus functional $J(\cdot)$, the computed error representation formula $\sum_{\kappa \in \mathscr{T}} \hat{\eta}_\kappa$, and the effectivity index $\theta = \sum_{\kappa \in \mathscr{T}} \hat{\eta}_\kappa / (J(\mathbf{u}) - J(\mathbf{u}_h))$. As noted in [23], here we see that, even on very coarse finite element meshes, the quality of the computed error representation formula is relatively good, in the sense that the effectivity indices are not too far away from unity. In particular, accuracy which

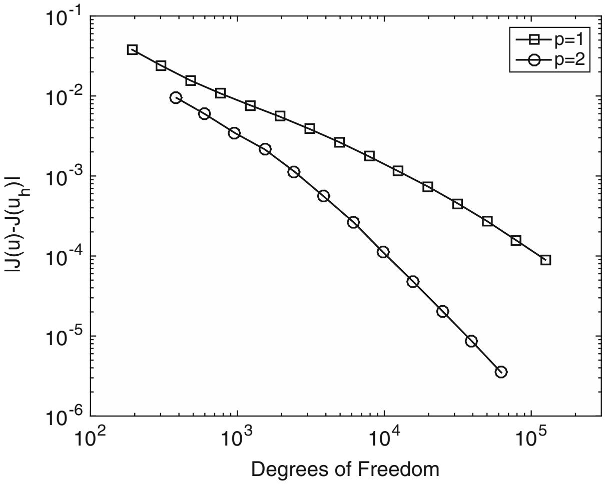


Fig. 3 Example 1: Convergence of the adaptive mesh refinement algorithm

is sufficient for practical/engineering calculations can be attained with a relatively small number of degrees of freedom. The results from Tables 1 and 2 are also shown in Fig. 3; here, we clearly observe the superiority of employing higher-order elements, in the sense that the error computed with $p = 2$ is significantly smaller than the corresponding quantity evaluated for $p = 1$, when the same number of degrees of freedom are employed.

Finally, in Fig. 4 we show the agglomerated polytopic meshes generated after 6 and 11 adaptive refinement steps have been performed with $p = 1$. Here we observe that the meshes have been refined in the vicinity of the reentrant corners present in the cross domain Ω as we would expect. Indeed, the regions away from these features have remained largely unrefined. Moreover, we note that hanging nodes are naturally generated through the application of local agglomeration-based refinement; this is very easily handled within the DG setting.

5.2 *Example 2: Modelling Trabecular Bone*

Following the work presented in the recent article [35], in this section we consider the evaluation of the (scaled) effective Young's modulus of the section of trabecular bone depicted in Fig. 1. The geometry Ω represents a cuboidal section of trabecular bone obtained by X-ray μCT scanning of a bone biopsy reconstructed from two-dimensional slices, cf. [31]. In this section we set the Young's modulus $E = 10\,\text{GPa}$ and the Poisson ratio $\nu = 0.3$; in this case, we computed the approximate reference value $J(\mathbf{u}) \approx 0.1236$. The initial fine mesh $\mathscr{T}_{\text{fine}}$, cf. Fig. 5a, consists of 1,179,569 tetrahedral elements, which is then agglomerated to generate a coarse polytopic

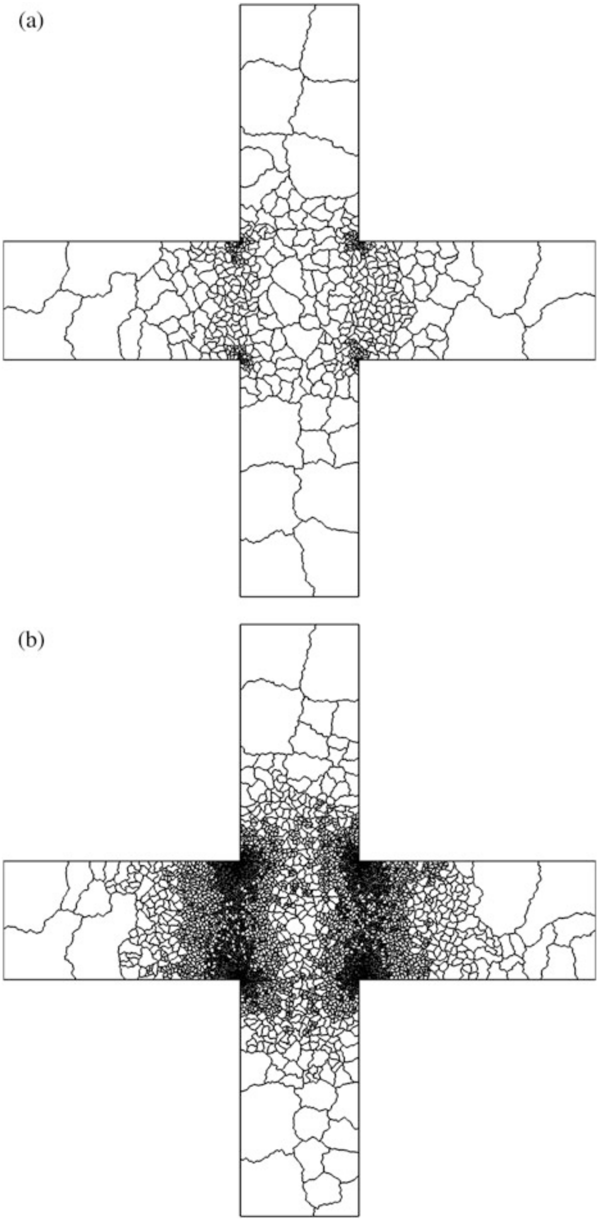


Fig. 4 Example 1. Agglomerated mesh for $p = 1$ after: (**a**) 6 refinements, with 514 elements; (**b**) 11 refinements, with 5274 elements

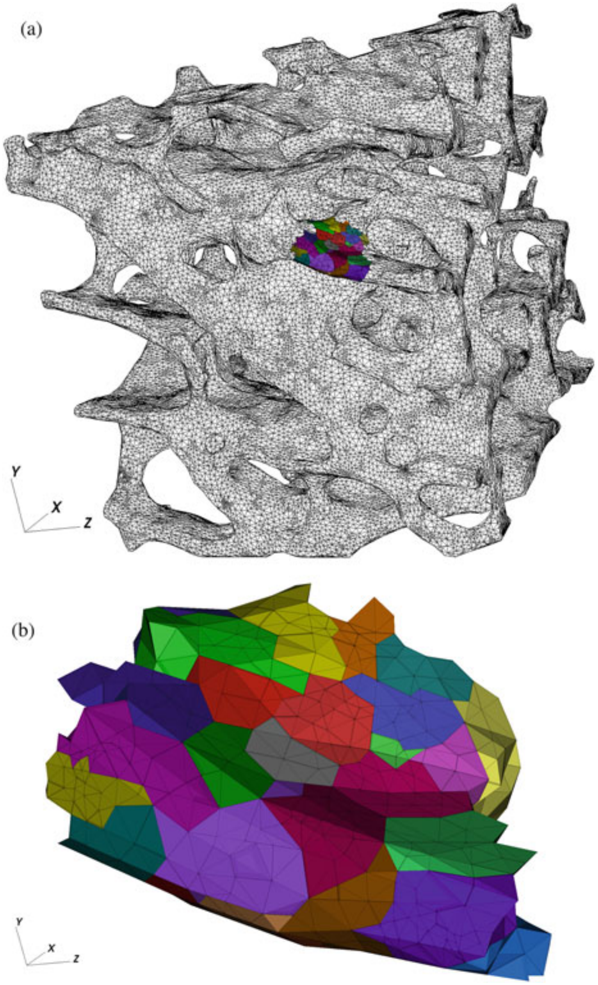


Fig. 5 Example 2. (**a**) Initial fine mesh; (**b**) Zoom of the first 30 agglomerated elements (shown in *colour*)

mesh $\mathscr{T}$ comprising of only 8000 elements. The first 30 elements are depicted in Fig. 5b. In Tables 3 and 4 we tabulate the results of the proposed adaptive refinement strategy with $p = 1$ and $p = 2$, respectively, as before, with $\texttt{REF} = 10\,\%$; cf., also, Fig. 6. As in the previous numerical experiment, we again notice that the effectivity indices θ are relatively good, given the coarse nature of the finite element meshes employed. Indeed, as the mesh is refined, we observe that θ improves and approaches unity. Again, here we observe that a sufficiently accurate (in terms of engineering accuracy) approximation to the target functional of interest may be computed with very few degrees of freedom. Finally, in Figs. 7 and 8 we show the primal and dual displacements, respectively.

Table 3 Example 2. Adaptive algorithm for $p = 1$

No of Eles	No of Dofs	$J(\mathbf{u}) - J(\mathbf{u}_h)$	$\sum_{\kappa\in\mathscr{T}} \hat{\eta}_\kappa$	θ
8000	96,000	−1.402E−01	−1.171E−01	0.83
13,600	163,200	−1.185E−01	−1.023E−01	0.86
22,994	275,928	−9.759E−02	−8.748E−02	0.90
38,867	466,404	−7.945E−02	−7.310E−02	0.92
65,634	787,608	−6.561E−02	−6.203E−02	0.95
110,752	1,329,024	−5.497E−02	−5.267E−02	0.96
18,6586	2,239,032	−4.452E−02	−4.374E−02	0.98
314,088	3,769,056	−3.517E−02	−3.585E−02	1.02

Table 4 Example 2. Adaptive algorithm for $p = 2$

No of Eles	No of Dofs	$J(\mathbf{u}) - J(\mathbf{u}_h)$	$\sum_{\kappa\in\mathscr{T}} \hat{\eta}_\kappa$	θ
8000	240,000	−6.009E−02	−3.872E−02	0.64
13,600	408,000	−4.153E−02	−3.065E−02	0.74
22,962	688,860	−2.473E−02	−2.106E−02	0.85
38,808	1,164,240	−1.660E−02	−1.606E−02	0.97
65,584	1,967,520	−1.253E−02	−1.253E−02	1.00
110,602	3,318,060	−8.375E−03	[a]	[a]

[a] Indicates that the dual problem was not computed

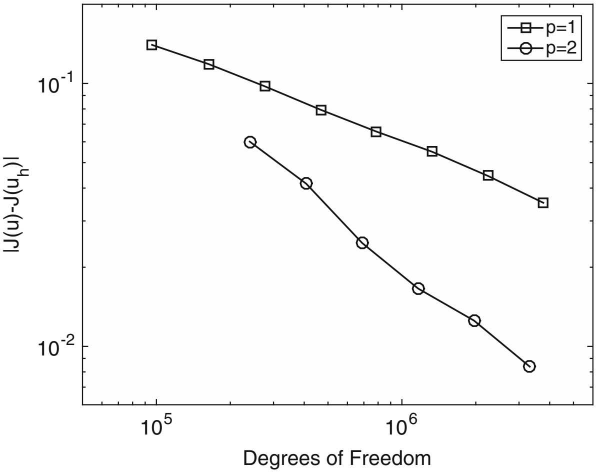


Fig. 6 Example 2: Convergence of the adaptive mesh refinement algorithm

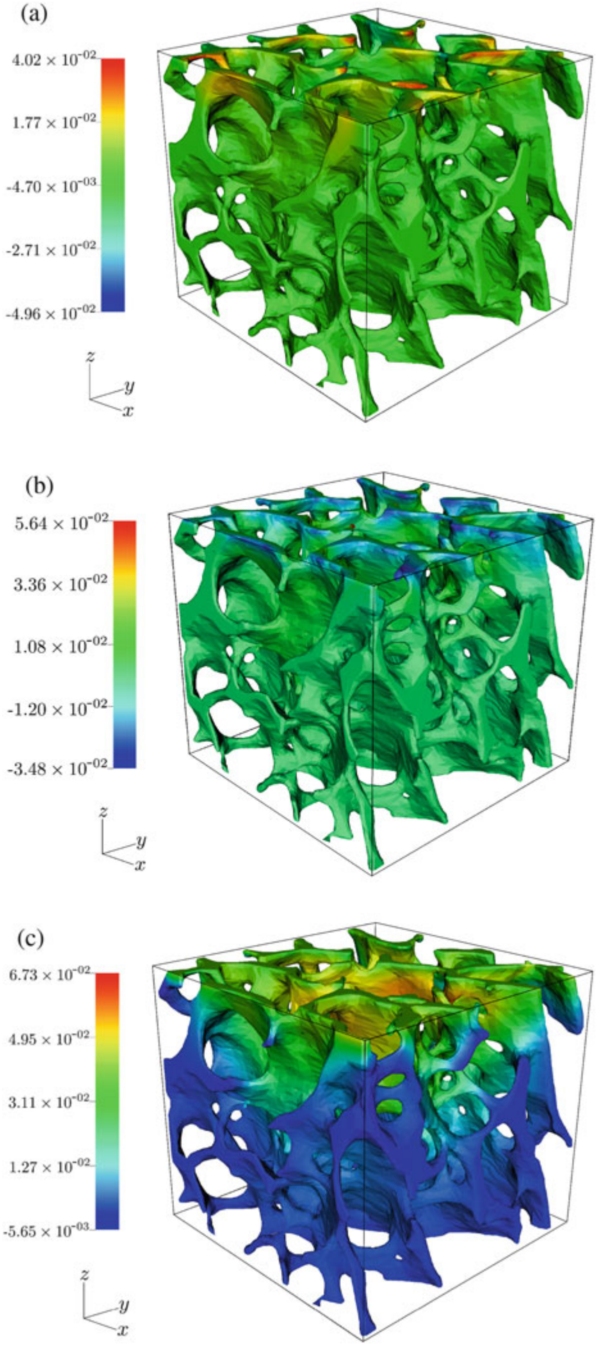


Fig. 7 Example 2. Approximate primal solution: (**a**) u_1; (**b**) u_2; (**c**) u_3

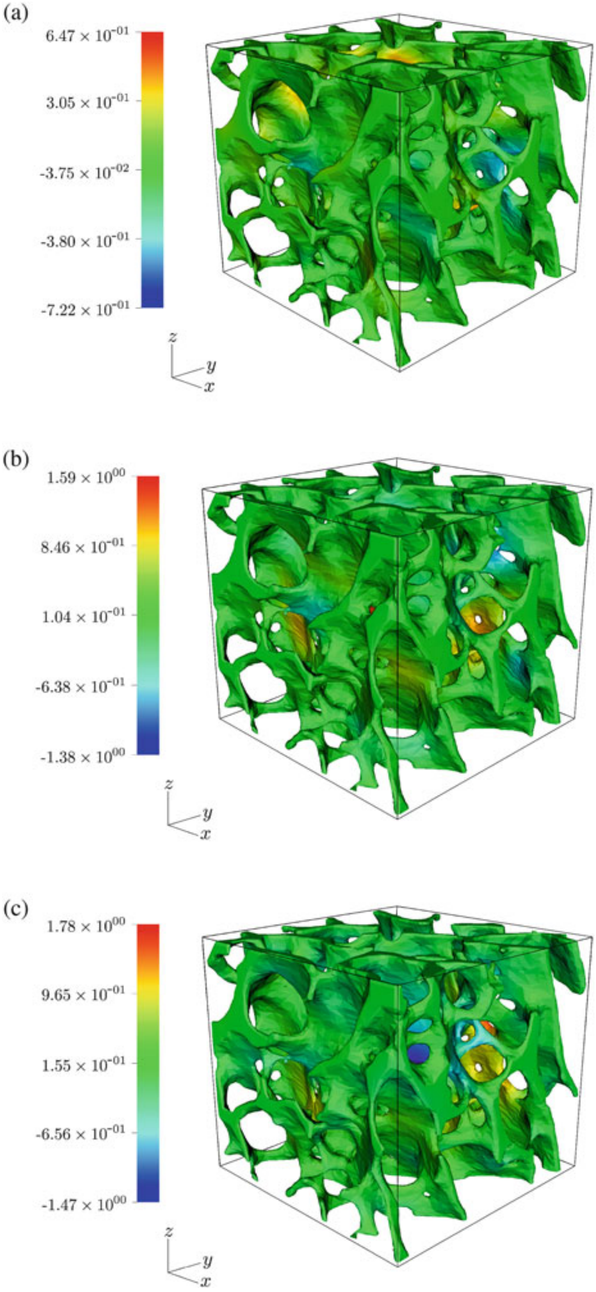


Fig. 8 Example 2. Approximate dual solution: (**a**) z_1; (**b**) z_2; (**c**) z_3

6 Conclusions

In this article we have developed an agglomeration-based adaptive mesh refinement algorithm within the goal-oriented setting for the DG approximation of the linear elasticity equations for a homogeneous isotropic material. We stress that this PDE model was chosen simply as a prototype problem; indeed, the work undertaken in this article naturally generalises to general classes of PDEs. By exploiting general agglomeration techniques, complicated geometries can be meshed using only a very small number of polytopic elements. The proposed adaptive refinement strategy then applies graph partitioning techniques to local elements which are marked for refinement. In this manner, the underlying mesh is adaptively refined on the basis of solution accuracy, rather than computing on fine meshes generated for the purposes of resolving the underlying geometry. This approach has been applied to a simple two-dimensional problem, as well as a test case involving estimating the (scaled) effective Young's modulus for a section of trabecular bone. Extensions of this work to include, for example, automatic *hp*-refinement will be undertaken as part of our programme of future research.

Acknowledgements Joe Collis acknowledges the financial support of the EPSRC under the grant EP/K039342/1.

References

1. Antonietti, P.F., Giani, S., Houston, P.: *hp*-Version composite discontinuous Galerkin methods for elliptic problems on complicated domains. SIAM J. Sci. Comput. **35**(3), A1417–A1439 (2013)
2. Antonietti, P.F., Giani, S., Houston, P.: Domain decomposition preconditioners for discontinuous Galerkin methods for elliptic problems on complicated domains. J. Sci. Comput. **60**(1), 203–227 (2014)
3. Antonietti, P.F., Cangiani, A., Collis, J., Dong, Z., Georgoulis, E.H., Giani, S., Houston, P.: Review of discontinuous Galerkin finite element methods for partial differential equations on complicated domains. In: Barrenechea, G.R., Brezzi, F., Cangiani, A., Georgoulis, E.H. (eds.) Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Lecture Notes in Computational Science and Engineering. Springer, Berlin (2015)
4. Antonietti, P.F., Houston, P., Sarti, M., Verani, M.: Multigrid algorithms for *hp*-version interior penalty discontinuous Galerkin methods on polygonal and polyhedral meshes (2015, submitted for publication)
5. Bassi, F., Botti, L., Colombo, A., Rebay, S.: Agglomeration based discontinuous Galerkin discretization of the Euler and Navier-Stokes equations. Comput. Fluids **61**, 77–85 (2012)
6. Bassi, F., Botti, L., Colombo, A., Di Pietro, D.A., Tesini, P.: On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations. J. Comput. Phys. **231**(1), 45–65 (2012)
7. Bassi, F., Botti, L., Colombo, A.: Agglomeration-based physical frame dG discretizations: an attempt to be mesh free. Math. Models Methods Appl. Sci. **24**(8), 1495–1539 (2014)

8. Bassi, F., Botti, L., Colombo, A.: h-Multigrid agglomeration based solver strategies for BR2 discontinuous Galerkin discretizations of elliptic problems (2015, Submitted for publication)
9. Becker, R., Rannacher, R.: An optimal control approach to a-posteriori error estimation in finite element methods. In: Iserles, A. (ed.) Acta Numerica, pp. 1–102. Cambridge University Press, Cambridge (2001)
10. Beirão da Veiga, L., Droniou, J., Manzini, G.: A unified approach for handling convection terms in finite volumes and mimetic discretization methods for elliptic problems. IMA J. Numer. Anal. **31**(4), 1357–1401 (2011)
11. Beirão da Veiga, L., Brezzi, F., Cangiani, A., Manzini, G., Marini, L.D., Russo, A.: Basic principles of virtual element methods. Math. Models Methods Appl. Sci. **23**(1), 199–214 (2013)
12. Brezzi, F., Buffa, A., Lipnikov, K.: Mimetic finite differences for elliptic problems. ESAIM Math. Modell. Numer. Anal. **43**(2), 277–295 (2009)
13. Burman, E., Hansbo, P.: An interior-penalty-stabilized Lagrange multiplier method for the finite-element solution of elliptic interface problems. IMA J. Numer. Anal. **30**, 870–885 (2010)
14. Burman, P., Hansbo, P.: Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method. Comput. Methods Appl. Mech. Eng. **199**, 2680–2686 (2010)
15. Burman, E., Hansbo, P.: Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method. Appl. Numer. Math. **62**, 328–341 (2012)
16. Cangiani, A., Manzini, G., Russo, A.: Convergence analysis of the mimetic finite difference method for elliptic problems. SIAM J. Numer. Anal. **47**(4), 2612–2637 (2009)
17. Cangiani, A., Georgoulis, E.H., Houston, P.: *hp*-Version discontinuous Galerkin methods on polygonal and polyhedral meshes. Math. Models Methods Appl. Sci. **24**(10), 2009–2041 (2014)
18. Cangiani, A., Dong, Z., Georgoulis, E.H., Houston, P.: *hp*-Version discontinuous Galerkin methods for advection–diffusion–reaction problems on polytopic meshes. ESAIM Math. Modell. Numer. Anal. **50**(3), 699–725 (2016)
19. Di Pietro, D.A., Ern, A.: Hybrid High-Order methods for variable diffusion problems on general meshes. C.R. Math. **353**, 31–34 (2014)
20. Di Pietro, D.A., Ern, A., Lemaire, S.: A review of Hybrid High-Order methods: formulations, computational aspects, comparison with other methods. In: Barrenechea, G.R., Brezzi, F., Cangiani, A., Georgoulis, E.H. (eds.) Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Lecture Notes in Computational Science and Engineering. Springer, Berlin (2015)
21. Fries, T.-P., Belytschko, T.: The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Methods Eng. **84**(3), 253–304 (2010)
22. Giani, S., Houston, P.: Domain decomposition preconditioners for discontinuous Galerkin discretizations of compressible fluid flows. Numer. Math. Theory Methods Appl. **7**(2), 123–128 (2014)
23. Giani, S., Houston, P.: Goal-oriented adaptive composite discontinuous Galerkin methods for incompressible flows. J. Comput. Appl. Math. **270**, 32–42 (2014)
24. Giani, S., Houston, P.: *hp*-Adaptive composite discontinuous Galerkin methods for elliptic problems on complicated domains. Numer. Methods Part. Diff. Equ. **30**(4), 1342–1367 (2014)
25. Hackbusch, W., Sauter, S.A.: Composite finite elements for problems containing small geometric details. Part II: implementation and numerical results. Comput. Visual Sci. **1**, 15–25 (1997)
26. Hackbusch, W., Sauter, S.A.: Composite finite elements for the approximation of PDEs on domains with complicated micro-structures. Numer. Math. **75**, 447–472 (1997)
27. Hansbo, P., Larson, M.G.: Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method. Comput. Methods Appl. Mech. Eng. **191**(17–18), 1895–1908 (2002)
28. Johansson, A., Larson, M.G.: A high order discontinuous Galerkin Nitsche method for elliptic problems with fictitious boundary. Numer. Math. **123**(4), 607–628 (2013)

29. Karypis, G., Kumar, V.: A fast and highly quality multilevel scheme for partitioning irregular graphs. SIAM J. Sci. Comput. **20**(1), 359–392 (1999)
30. Massing, A.: Analysis and implementation of finite element methods on overlapping and fictitious domains. Ph.D. thesis, University of Oslo (2012)
31. Perilli, E., Baruffaldi, F.: Proposal for shared collections of X-ray microCT datasets of bone specimens. In: ICCB03, Zaragoza, Spain, 24–26 September 2003
32. Shewchuk, J.R.: Triangle: engineering a 2D quality mesh generator and Delaunay triangulator. In: Lin, M.C., Manocha, D. (eds). Applied Computational Geometry: Towards Geometric Engineering. Lecture Notes in Computer Science, vol. 1148, pp. 203–222. Springer, Berlin (1996); From the First ACM Workshop on Applied Computational Geometry
33. Si, H.: TetGen, a Delaunay-based quality tetrahedral mesh generator. ACM Trans. Math. Softw. **41**(2), 11:1–11:36 (2015)
34. Sukumar, N., Tabarraei, A.: Conforming polygonal finite elements. Int. J. Numer. Methods Eng. **61**(12), 2045–2066 (2004)
35. Verhoosel, C.V., van Zwieten, G.J., van Rietbergen, B., de Borst, R.: Image-based goal-oriented adaptive isogeometric analysis with application to the micro-mechanical modeling of trabecular bone. Comput. Methods Appl. Mech. Eng. **284**, 138–164 (2015)
36. Wihler, T.P.: Locking-free adaptive discontinuous Galerkin FEM for linear elasticity problems. Math. Comput. **75**(255), 1087–1102 (2006)

Part IV
Advances in XFEM/Fictitious Domain Methods

Stabilized X-FEM for Heaviside and Nonlinear Enrichments

Giulio Ventura and Claudia Tesei

Abstract One of the drawbacks of the eXtended Finite Element Method and similar approaches, like the Generalized Finite Element Method, is the problem of ill-conditioning of the related systems of equations at the solution stage. This occurs for example in Heaviside function enrichments when the discontinuity is close to discretisation nodes but also for non-linear enrichment functions used in conjunction to geometric enrichment domains. In the present work the motivation of ill-conditioning is analyzed to derive a novel methodology for stabilization, based on setting proper constraints for the variables. This methodology does not impact on the initial formulation nor in the element stiffness computation, so that it is very effective for engineering applications. Results are analyzed in 1D and 3D to show its performance and properties.

1 Introduction

The success of enriched formulations like XFEM or GFEM [2, 4, 6, 13, 16, 37] is nowadays well recognized also in commercial softwares [1, 23], implementing many features of these methods to the end of achieving more accurate solutions with coarser meshes and independency of discretization on the position of features of the problem like cracks or singularities. However, some fundamental problems are still present like quadrature at the element level [14, 24, 26, 27, 31, 32, 34, 35] and ill-conditioning of the resulting systems of equations. These problem are solved in practical applications by some techniques that require additional computational resources or introduce approximations of difficult determination.

The problem of ill-conditioning, although known since the beginning of the development of these methods, has found limited attention in literature. The main techniques used for stabilization can be divided into three groups:

- perturbation or preconditioning of the global system of equations. One of the first contribution based on a perturbation scheme was given by Strouboulis et al.

G. Ventura (✉) • C. Tesei
DISEG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy
e-mail: giulio.ventura@polito.it; claudia.tesei@polito.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_10

[30] in the context of GFEM. Approaches based on special preconditioners were presented by Béchet et al. [5], Menk and Bordas [25], Berger-Vergiat et al. [12], Waisman and Berger-Vergiat [36], Lang et al. [21] to cite some of the many works based on this technique;

- element stiffness matrix kernel space correction. This approach, proposed in [22], is based on analyzing the element stiffness matrix eigenspace and introduce a threshold and a stabilization factor. The element matrix is modified by adding, for each eigenvalue needing stabilization, a term given by the stabilization factor times the outer product of the corresponding eigenvector times itself. This method proves effective but its implementation is not straightforward;
- the Stable Generalized Finite Element Method (SGFEM), introduced by Babuška and Banerjee [3], and recently reformulated for higher order shape functions in [38]. This method is conceptually quite attractive and is based on eliminating linearly dependent parts in the approximation space by subtracting a polynomial interpolation of the enrichment function. An application of this method to 3D fracture has been given in [18].

Finally, a very recently introduced approach, called Orthonormalized Generalized Finite Element Method (OGFEM) has been introduced in [29]. In OGFEM the ill-conditioning is prevented by orthonormalizing the basis functions to eliminate linear dependence in the approximation basis.

The present paper proposes a new method for stabilization that is conceptually different from the literature approaches and is conceived for simplicity and effectiveness to have the least impact on implementation and computational cost. This will be accomplished by introducing suitable constraints on the enrichment variables. The outline of the paper follows: in Sect. 2 it is first examined the motivation for ill-conditioning; in Sect. 3 the analytic structure of ill conditioning is derived and commented. Then the new stabilization technique is introduced in Sect. 4 and the method is validated by analyzing 1D and 3D results in Sect. 5.

2 Motivation for Ill-Conditioning

Reference will be made in the following to linear elastic solid mechanics problems, although the discussion can be applied to general enriched finite element discretizations. Let $\mathbf{u}$ be the displacement field, ψ an enrichment function and $\mathbf{N}$ the elements shape functions. The dependence of the functions w.r.t. the position in space will be omitted for the sake of brevity.

The standard formulation in XFEM/GFEM is:

$$\mathbf{u} = \sum_i \mathbf{N}_i \mathbf{u}_i + \sum_i \psi \mathbf{N}_i \mathbf{a}_i \tag{1}$$

where the first addend is the standard finite element approximation and the second adds to the approximation space the enrichment function ψ by partition of unity. Elements where the enrichment function is not needed will miss the second addend. Transition from enriched to non-enriched elements may need blending techniques as illustrated in [15, 34]. The symbols in (1) are defined as follows:

- the sum on i is extended to the nodes of the finite element where the evaluation point of the displacement function $\mathbf{u}$ is located;
- $\mathbf{u_i}$ are the classical finite element displacement nodal variables;
- $\mathbf{a_i}$ are the enrichment nodal variables.

When the approximation (1) is introduced in the variational principle stating the mechanical problem, whenever the second addend will have terms linearly dependent on terms in the first addend then ill-conditioning of the resulting system of equations will occur. This ill-conditioning expresses the fact that the same solution can be represented by linearly dependent sets of variables $\mathbf{u}_i$ and $\mathbf{a}_i$ (indeterminate system of equations). An analytical example will be presented in the next Section.

This situation may occur in several cases. The most common and intuitive is when the generalized Heaviside step function H is used as enrichment function for representing jumps in displacement (e.g. cracks). Let d be the signed distance from the discontinuity, it is

$$H(d) = \begin{cases} +1 \text{ for } d \geq 0 \\ -1 \text{ for } d < 0 \end{cases} . \tag{2}$$

The generalized Heaviside being a piecewise constant function, it will be linearly independent of the element shape functions whenever the discontinuity surface $d = 0$ will cross inside an element. Strict linear dependence will hold when the discontinuity will cross a node and numerical ill-conditioning is observed as the discontinuity will get close to a node. This problem is sometimes solved by moving the crack trajectory away from nodes or by moving the nodes away from the crack surface with the consequence of introducing an a priori unknown source of further approximation [22]. However, not only ill-conditioning takes place with piecewise constant or linear enrichment functions (spanned by element shape functions). Ill-conditioning also arises when employing non-linear enrichments in conjunction with the concept of geometric enrichment area.

The concept of geometric enrichment was first introduced by Ventura et al. in [33] and then, independently and almost simultaneously, by Laborde et al. [20] and Béchet et al. [5]. It is recalled that by geometric enrichment it is meant a fixed domain area where finite elements are enriched (independently of the mesh size), while in topologic enrichment certain specific elements are enriched so that, when the mesh size is reduced, the enrichment area shrinks accordingly.

It was observed, in fact, that if a near crack tip enrichment (or dislocation core enrichment) is confined to the element containing the crack tip (or dislocation core), as the discretization mesh is refined the solution error in energy increases instead of decreasing. On the other hand, if a fixed domain area is enriched, the error decreases by decreasing the mesh size, as expected [33].

We can explain this behaviour by considering that a singularity is present at the tip, so that the solution is highly nonlinear in a neighbour of the tip and tends to flatten away. The enrichment function depicts the solution closely, so that the element shape functions play the role of adjusting the enrichment to the particular case at hand. If, refining the mesh, the enriched domain is shrunk, then the high gradients and nonlinearity of the solution must be represented by polynomials (the representation space of standard finite elements), with the consequent degradation of the solution approximation. In contrast, this will not occur for geometric enrichment, being the enriched domain independent of the mesh size.

To see how this may affect ill-conditioning, consider the graph of the enrichment function for a dislocation in uniform shear stress considered in [33], Fig. 1. The function has been analytically derived by Koehler [19] and other reference solutions can be found in [28]. It is immediate to observe that, while the enrichment function is markedly nonlinear close to the dislocation core, its graph tends to a linear function

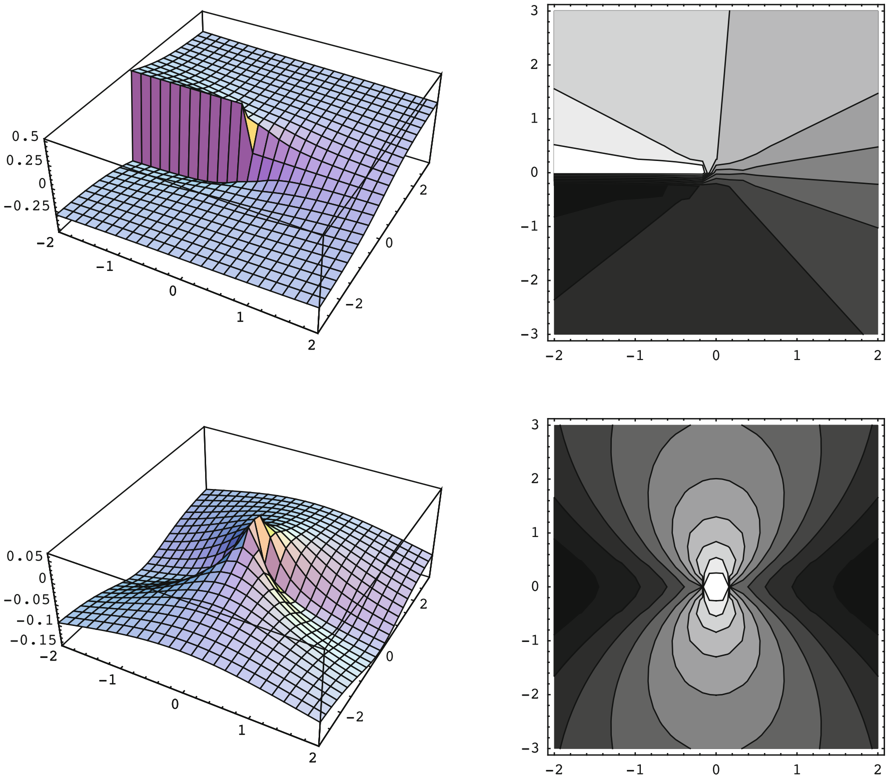


Fig. 1 Graph of enrichment functions for edge dislocation problems. The two graphs represent the displacement components of the analytic solution [19]

as the distance from the core is increased. Consequently, there will exist a size of the geometric enrichment for which the enrichment function will be *numerically* linear and the global system of equations will be indeterminate. This consideration is the keystone of SGFEM [3], where the linear interpolation to the enrichment function is subtracted from the enrichment space. The same reasoning holds for other enrichments like near field functions for cracks.

3 Structure of Ill-Conditioning

The new method for stabilization proposed in the present paper is based on the observation of the analytic structure of the problem in the simple case of a bar in two cases:

- Heaviside function enrichment, representing the physical problem of a cut separating the bar into two parts. This is typical of crack faces enrichment;
- linear function enrichment, representing the limit behaviour of a nonlinear function typical, for example, of a near crack tip enrichment or a near dislocation core enrichment when far from the tip or core, respectively.

It is interesting to note that ill-conditioning is also observed, for the reasons stated above, when regularized Heaviside functions are directly used as enrichment functions according to (1). This, however, does not occur when a proper variational formulation is adopted, like in [10]. The formulation introduced in [10] and employed also in [7–9, 11] is in fact numerically stable for any value of the Heaviside function regularization parameter and size of the enrichment domain.

3.1 *Heaviside Function Enrichment*

Consider a bar with a cut placed at a distance d from the left end, subjected to prescribed displacements δ_1 and δ_2 at the left and right ends, respectively, Fig. 2. To get more readable results, unit bar length, cross section area and material Young's modulus are assumed. The solution of this problem is given by zero strain and stresses and rigid body motions of the two parts separated by the cut.

Consider a discretisation with one finite element and let

$$N_1(x) = 1 - x \quad ; \quad N_2(x) = x \tag{3a}$$

$$\psi(x) = H(x) = \begin{cases} +1 & \text{for}\, x \geq d \\ -1 & \text{for}\, x < d \end{cases} \tag{3b}$$

$$u(x) = \sum_{i=1}^{2} N_i u_i + \sum_{i=1}^{2} \psi(x) N_i a_i \tag{3c}$$

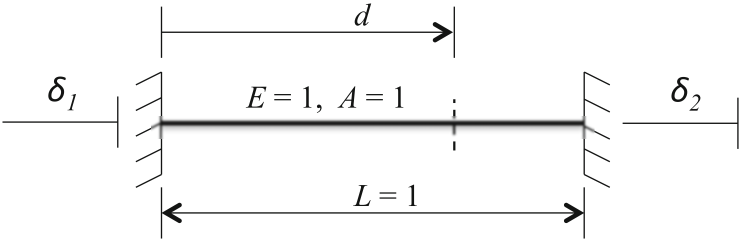


Fig. 2 A bar with a cut discontinuity and prescribed ends displacement

where the subscript 1 stands for the left end and the subscript 2 for the right end. Enforcing the boundary conditions by Lagrangian multipliers r_1, r_2, the total potential energy functional of the system is

$$\Pi = \frac{1}{2}\int_0^L A\sigma(x)\varepsilon(x)\,\mathrm{d}x + r_1\,(u(0) - \delta_1)) + r_2\,(u(L) - \delta_2)) \tag{4}$$

with the strain and stresses being given by, respectively

$$\varepsilon(x) = \frac{\mathrm{d}u(x)}{\mathrm{d}x} \quad ; \quad \sigma(x) = E\varepsilon(x). \tag{5}$$

The stationarity conditions w.r.t. the variables $(u_1, u_2, a_1, a_2, r_1, r_2)$ yield the following solutions, depending on the position of the discontinuity

$$\begin{array}{lll} r_1 = r_2 = 0 & \forall d \text{ i.e. constraint reactions = 0 as expected} & \\ d \neq 0 \text{ and } d \neq L \rightarrow & u_1 = u_2 = \frac{\delta_1+\delta_2}{2};\ a_1 = a_2 = \frac{\delta_2-\delta_1}{2} & \\ d = 0 & \rightarrow u_2 + a_2 - \delta_2 = 0; & u_1 = \frac{\delta_1+\delta_2}{2};\ a_1 = \frac{\delta_2-\delta_1}{2} \\ d = L & \rightarrow u_1 - a_1 - \delta_1 = 0; & u_2 = \frac{\delta_1+\delta_2}{2};\ a_2 = \frac{\delta_2-\delta_1}{2}. \end{array} \tag{6}$$

The solution (6) is unique for $d \neq 0$ and $d \neq L$ (discontinuity along the bar) and is indeterminate for $d = 0$ and $d = L$ (discontinuity at the ends of the bar). In particular, when the solution is indeterminate, a standard variable is linearly dependent on the corresponding enrichment variable, yielding ∞^1 solutions. Being $H(0) = -1$, $H(L) = +1$, (3b), the general structure of the indeterminate equation is

$$u_i + H(x_i)a_i - \delta_i = 0. \tag{7}$$

Equation (7) is derived from (6) looking at the cases $d = 0$ and $d = L$, first equation after the right arrows.

3.2 Linear Function Enrichment

Consider a bar subjected to prescribed displacements δ_1 and δ_2 at the left and right ends, respectively, Fig. 3. Again, to get more readable results an unit bar length, cross section area and material Young's modulus are assumed. Consider the following linear enrichment function

$$\psi(x) = c_0 + c_1 x. \tag{8}$$

The solution of this problem is given by linearly varying displacement from δ_1 to δ_2 and constant strain $(\delta_2 - \delta_1)/L$ as if the enrichment function would not be present. It is useful to mark that this enrichment function has no physical sense, but it represents a decaying nonlinear enrichment function near the boundary of a fixed geometric enrichment area.

For better clarity, the cases $c_0 = c_1 = 0$ (no enrichment), $c_0 \neq 0$, $c_1 = 0$ (constant enrichment), $c_0 = 0$, $c_1 \neq 0$ (linear homogeneous enrichment) and $c_0 \neq 0, c_1 \neq 0$ (general linear enrichment) will be examined.

Following the same steps as in the previous section, replacing the expression for $\psi(x)$ with (8), the stationarity conditions yield the following solutions in the different cases

$$\begin{array}{lcl}
r_1 = \delta_2 - \delta_1,\ r_2 = \delta_1 - \delta_2\ \forall c_0,\ c_1 & & \text{constraint reactions} \\
c_0 \neq 0,\ c_1 = 0 & \rightarrow & u_1 + c_0 a_1 - \delta_1 = 0;\ u_2 + c_0 a_2 - \delta_2 = 0 \\
c_0 = 0,\ c_1 \neq 0 & \rightarrow & u_1 = \delta_1;\ u_2 + a_2 c_1 - \delta_2 = 0;\ a_2 = a_1 \\
c_0 \neq 0,\ c_1 \neq 0 & \rightarrow & c_0 u_2 - \psi(L)(u_1 - \delta_1) - c_0 \delta_2 = 0; \\
 & & u_1 + c_0 a_1 - \delta_1 = 0;\ u_1 + c_0 a_2 - \delta_1 = 0.
\end{array} \tag{9}$$

As in the case of the Heaviside function enrichment, it is immediate to observe that the indeterminate equations tie standard and enrichment variables u_i and a_i. Consequently, when ill conditioning occurs, it is generated by the fact that infinite sets of variables u_i and a_i are solution to the linear system. Then, adding proper

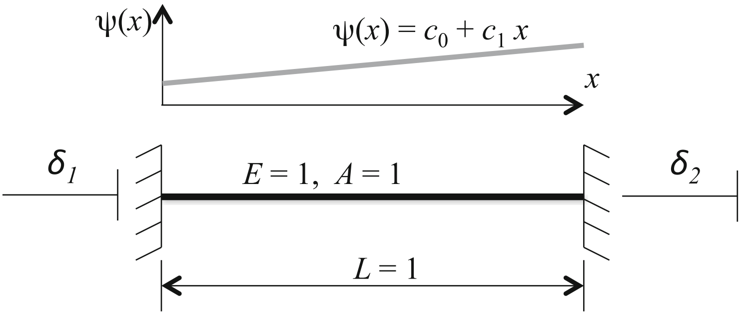


Fig. 3 A bar with a linear enrichment function and prescribed ends displacement

additional constraint equations for the variables can resolve indetermination, as shown in the next Section.

4 The New Stabilization Technique

To eliminate the indeterminacy a fictitious constraints can be introduced for the variables. Starting from the study of Eqs. (6) and (9) several forms of additional constraints have been studied. For example, in the case of Heaviside function enrichment, considering (7), an additional constraint equation of the form

$$u_i - H(x_i)a_i = 0 \tag{10}$$

has been tested and allows to obtain good results. The rationale in (10) is that it is an equation always linear independent from (7), as $H(x_i)$ shows reversed sign in the two equations. Although (10) works nicely with Heaviside enrichments, its analytical study shows that it is not able of eliminating indeterminacy in the case of nearly linear enrichments.

The most efficient and universal stabilization equation that has been found is simply biasing to zero the enrichment variables a_i. The term "biasing" and not "setting" is used as, when the indeterminacy holds strictly (i.e. the Heaviside function has the discontinuity at a node or the additional enrichment function is linear), then the enrichment variables a_i can be set to zero. However, when indeterminacy is caused by nearly linear enrichment or by an Heaviside function having the discontinuity in proximity of a node, setting the enrichment variables to zero will cause approximations of difficult determination.

The ideal stabilization should be such that, if indeterminacy would take place, it is stabilized. If applied to elements and cases where it is not needed it will leave the solution unaltered.

Therefore, the enrichment variables a_i are biased to zero by adding a quadratic penalty term to the total potential energy functional as follows

$$\Pi = \frac{1}{2}\int_0^L A\sigma(x)\varepsilon(x)\,\mathrm{d}x + r_1\left(u(0)-\delta_1\right)\big) + r_2\left(u(L)-\delta_2\right)\big) + \frac{1}{2}\alpha\sum_{i=1}^{2} a_i^2. \tag{11}$$

The added term states, in the penalty method light, that the enrichment variables a_i should be zero for each i. Enforcing strictly this condition would produce an erroneous solution to our problem, and this would happen for high values of the penalty parameter α. However, as the indeterminacy of the system of equations is to be ruled out, the minimum possible value for α will be chosen to recover a numerically well conditioned problem.

Therefore, in contrast to penalty methods, where the penalty parameter is set very high to enforce as more precisely as possible the constraint (leading again to ill conditioning), here the constraint is fictitious and the penalty parameter is set the lowest possible. Physically, this is equivalent to introduce a kind of weak spring for the enriched variables.

Note that, as the penalty term is not integrated at the element level, the stabilisation can be applied on the global stiffness matrix and its implementation is straightforward in any dimensions and for any element type.

5 Method Validation

The proposed methodology is applied to some 1D to 3D examples to show its effectiveness. In particular, the 1D problems of the bar with a cut and the bar with linear enrichment will be examined. Then, two 3D examples of a parallelepiped with a crossing surface discontinuity will be shown to observe the marked improvement in problem conditioning and the limited approximation introduced by the proposed technique.

5.1 1D Bar with a Cut

Reference is made to the 1D problem of the bar with a cut of Fig. 2. Although in this case stabilization is needed only when d approaches 0 or L, it is useful to observe the influence of the stabilization term in the solution when the discontinuity is placed far from the ends. If the stabilization will not alter the solution even when not needed, the problem of distinguishing whether it is needed or not will be automatically ruled out, and this is an important objective of the present technique in order to drastically simplify its application. Of course, criteria can be introduced to selectively apply stabilization and reduce stabilization induced approximation.

The exact solution of the problem of Fig. 2 for $\delta_1 = 0$, $\delta_2 = 1$ is given by zero strain, zero displacement on the left of the cut ($x < d$), unit displacement on the right of the cut ($x > d$). Figure 4 shows in the three columns the solution for $d = 0$, $d = 0.3\,L$, $d = L$. In the first two rows the plots of displacement and strain are given for a proper value of the penalty parameter. In the present case, numerical experiments have shown that a good value for the penalty parameter α is given by $1/1000$ the maximum diagonal term in the stiffness matrix, giving a system condition number (intended as ratio of the largest to the smallest matrix eigenvalue) of 4250 when the discontinuity is at one end of the bar.

Finally, the last row in Fig. 4 shows plots of the condition number with α. It can be observed that, when stabilization is not needed ($d = 0.3\,L$), any value of alpha does not impact significantly the condition number. On the other hand, when stabilization is essential to get the problem solution by ruling out indeterminacy, excessive low values of the penalty parameter cause the condition number to blow up as the stabilization term tends to disappear.

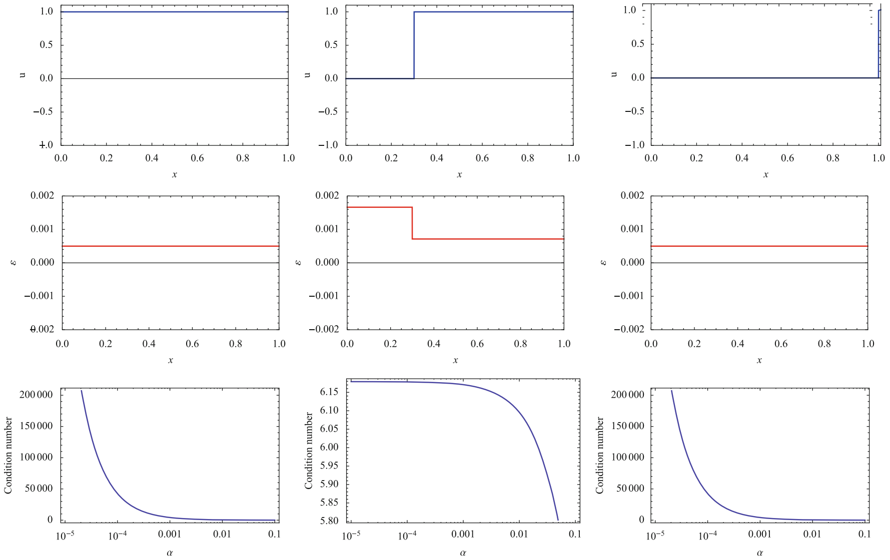


Fig. 4 Stabilized solutions for the reference problem of the bar with a cut of Fig. 2. Columns show three positions of the discontinuity ($d = 0,\ 0.3L,\ L$). The first two rows show displacement and strain for $\alpha = 0.001$ while the last row plots the system condition number with varying α

The solutions of Fig. 4 show displacement practically equal to the exact solution and the approximation introduced by the present stabilization technique is visible in the strain plot where, instead of zero strains (exact solution), a small error is observed, being ε below 2‰. In the cases where stabilization is really needed (first and third column of Fig. 4) the error in strain is 0.5‰. This approximation is quite good if it is considered that the strain in absence of discontinuity is $\delta_2/L = 1$, so that the error in strain introduced by the stabilization is of the order of 1‰.

To conclude the study of the present example, plots of the solution in the case $d = 0$ (cut at the left end) are given for three values of α in Fig. 5. Solution in terms of displacement and strain are plotted for $\alpha = 1.0,\ 0.1,\ 0.01$. These values have been chosen large otherwise the effect in the approximation introduced by stabilization would not be visible. It can be observed how large values of the stabilization parameter give a slope to the displacement plot and larger residual strain.

5.2 1D Bar with Linear Enrichment

Reference is made to the 1D problem of the bar with a linear enrichment function (8) of Fig. 3 representing the degenerate case of a nonlinear enrichment function that has become approximately linear in the element. As previously observed, in this case

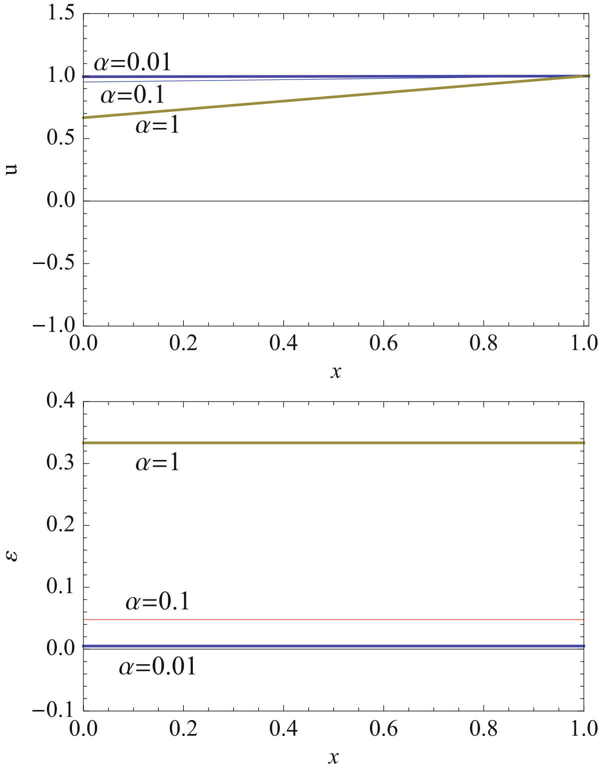


Fig. 5 Stabilized solutions for the reference problem of the bar with a cut of Fig. 2 for large values of the penalty parameter α

stabilization is needed for any values of the constants c_0 and c_1 as the enrichment function ψ is already spanned by the element shape functions (the enrichment function is linearly dependent on the shape functions).

The exact solution of this problem for $\delta_1 = 0$, $\delta_2 = 1$ is given by unit constant strain and linearly varying displacement from zero at the left end to one at the right end, for any values of the constants c_0 and c_1 defining the enrichment function.

Before showing the effectiveness of the method, it is useful to observe the structure of the linear system involved in this case to set the penalty stabilization parameter properly. For how the case has been constructed, the matrix of coefficients of this system is given by the enriched element stiffness matrix plus the two rows and columns of the Lagrangian multipliers enforcing the Dirichlet (prescribed displacement) boundary conditions. The system matrix is given by (12), where the system variables are ordered as follows: u_1, u_2, a_1, a_2, r_1, r_2.

$$\begin{bmatrix} 1 & -1 & c_0 & -c_0 - c_1 & 1 & 0 \\ -1 & 1 & -c_0 & c_0 + c_1 & 0 & 1 \\ c_0 & -c_0 & c_0^2 + \frac{c_1^2}{3} + 2\alpha & -c_0^2 - c_1 c_0 - \frac{c_1^2}{3} & c_0 & 0 \\ -c_0 - c_1 & c_0 + c_1 & -c_0^2 - c_1 c_0 - \frac{c_1^2}{3} & c_0^2 + 2c_1 c_0 + \frac{4c_1^2}{3} + 2\alpha & 0 & c_0 + c_1 \\ 1 & 0 & c_0 & 0 & 0 & 0 \\ 0 & 1 & 0 & c_0 + c_1 & 0 & 0 \end{bmatrix} \tag{12}$$

Looking at the diagonal terms of the enriched variables part (third and fourth row/column) they are proportional to the penalty stabilization parameter α and to the square of the coefficients c_0 and c_1, i.e. roughly to the square of the maximum enrichment function value in the element. Therefore, two considerations can be made:

- large enrichment function values would cause anyway ill-conditioning. However, this case is not of interest in applications as nonlinear enrichment functions usually assume vanishing values far from the features that are to be represented (for example crack tip enrichments far from the singularity);
- the value of the penalty stabilization parameter must be chosen to be suitably large compared to the diagonal stiffness terms otherwise its contribution will have no effect.

In the present example, in contrast to the previous case of the bar with a cut, the exact solution is always obtained for any value of the stabilization penalty parameter ranging by several orders of magnitude, Fig. 6. What changes is the system condition number, that in this case is influenced both by α and by c_0 and c_1.

This has been studied in the following Table 1, where the system condition number is reported for varying values of c_0 and c_1 in the range $1 \ldots 1000$ for a penalty parameter α given by $1/1000$ the maximum diagonal term in the stiffness matrix. When the enrichment coefficients become large compared to element stiffness terms, ill conditioning is observed anyway as previously pointed out, but this feature is independent of the proposed stabilization method.

5.3 *Three Dimensional Test*

The proposed stabilization technique is finally tested on two 3D benchmark examples, similar to the one dimensional cases where the problem has been studied analytically. However, while in the 1D cases the stabilization parameter has been fixed arbitrarily, here a study is presented for selecting the best stabilization penalty parameter in terms of optimal system conditioning. The two considered cases are:

- a bar with a slant cut with standard Heaviside function enrichment;
- a bar with an orthogonal cut where the enrichment is a regularized Heaviside function representing the case of quasi linear enrichment.

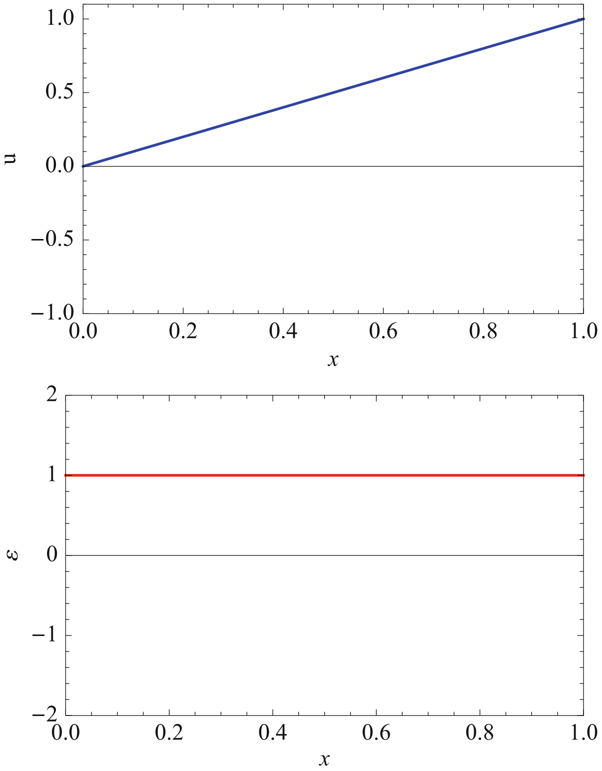


Fig. 6 Stabilized solutions for the reference problem of the bar with linear enrichment of Fig. 3

Table 1 Values of the system condition number for varying enrichment coefficients and constant value of the stabilization parameter

$\mathbf{c_0}$	0	1	10	1	100	1	1000	1	1000
$\mathbf{c_1}$	0	1	1	10	1	100	1	1000	1000
Condition number	1E3	3E3	1E5	3E4	1E7	3E6	1E9	3E8	2E9

The physical problem is similar to the example in Fig. 2. However, a 3D parallelepiped is used in place of the 1D bar element. The problems have been studied by the software X3D developed by the Author. It is a FORTRAN code using GMSH [17] as pre and post-processor. The code is able of dealing with problems of several hundred thousand elements with high efficiency and has been developed as research platform.

In the two examples a linear elastic material with a Young modulus $E = 10\,\text{GPa}$ and a Poisson coefficient $\nu = 0$ is assumed. The parallelepiped has one end fixed and the other displaced by 1 mm in the longitudinal direction. Let $\kappa = \max(K_{ii})$ the value of the largest diagonal element in the stiffness matrix. The penalty

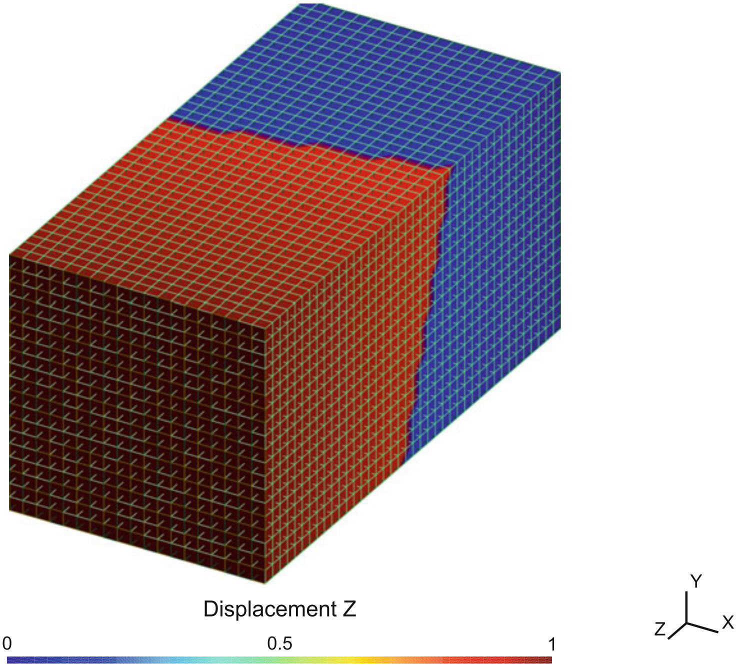


Fig. 7 A 3D parallelepiped with a slant cutting plane and prescribed displacements at the ends

stabilization parameter is set as

$$\alpha = \frac{\kappa}{\omega} \tag{13}$$

where ω is a scaling factor that will be varied to show how the condition number is affected by the stabilization. The largest the ω the smallest the stabilization penalty parameter will be set.

In the first case, a parallelepiped with dimensions $100 \times 100 \times 400$ mm is considered, with a mesh of 14,440 hexahedral elements ($19 \times 19 \times 40$ elements) and the discontinuity is a slanted cut. In Fig. 7 the displacement contours at solution show the physics of the problem.

This problem cannot run without stabilization as the linear system solver exits with zero pivot error detection and numerically infinite condition number unless a cutting plane avoiding all element nodes is considered.

To fix the order of magnitude of the condition number, the same problem, analyzed as linear elastic (without discontinuity), has a system condition number of 6E + 5.

Table 2 Values of the system condition number for varying stabilization parameter scaling factor in the 3D cut example

ω	1E + 3	1E + 4	1E + 5	1E + 6
Condition number	2E + 6	2E + 6	4E + 6	4E + 7

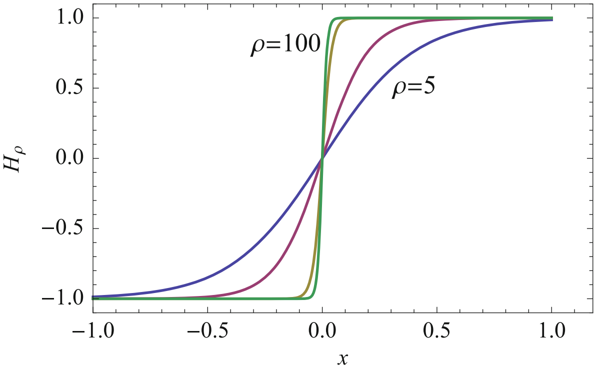


Fig. 8 Graph of the regularized Heaviside function (14)

Table 2 summarizes how the system condition number varies with the stabilization penalty parameter scaling factor ω when the cut discontinuity is introduced. As observed in the one-dimensional study, the smallest stabilization penalty parameter (and therefore the largest ω) will yield the lowest approximation error in terms of residual traction between the two parts separated by the cut. The best result is obtained for $\omega = 1\text{E} + 5$, that is the largest ω value not implying an increase in condition number. In this case the residual stress due to the introduction of the penalty stabilization has been found to be 0.02 MPa, that is 8‰ of the stress in the linear elastic case with no discontinuity (equal to 25 MPa).

This result is quite good and evidences the only limitation of the methodology, i.e. when very large displacements are present the closing residual stresses due to the penalty stabilization may be large. However, the methodology is certainly sound for small displacement mechanics.

The case of the bar with quasi linear enrichment is equal to the previous example but, instead of using an exact Heaviside function enrichment, its regularized form has been considered

$$H_\rho(d) = \frac{2}{\pi} \arctan(d/\rho) \qquad (14)$$

Function (14) is plotted in Fig. 8. For decreasing ρ values tends to the exact Heaviside function (2). Far from the discontinuity this function can be regarded as linear or approximately constant. When this function is inserted as enrichment function ψ into the XFEM displacement approximation (1) and an enrichment domain of width Δ extending over several elements is considered, the expected behavior is:

- a well conditioned problem when the enrichment domain width Δ is small compared to the regularization parameter ρ. This because for Δ comparable to ρ the function is markedly nonlinear in the enrichment domain;
- a progressively more ill conditioned problem as the enrichment domain width Δ increases because the enrichment function becomes approximately linear in the enriched elements.

It is important to mark again that the purpose of this example is to illustrate how ill-conditioning may arise when using enrichment functions that decay to linear functions with the distance from features and these functions are directly introduced in the expression (1). However, when a proper variational formulations is used [8, 10, 11], the ill-conditioning issue may be automatically ruled out.

This problem has been studied on a $100 \times 100 \times 1200$ mm parallelepiped discretized by $9 \times 9 \times 30$ elements (2430 hexahedra), same material properties and boundary conditions as the previous example and a value for the regularization parameter $\rho = 0.003$.

The expected behavior is fully confirmed by numerical experiments, whose results are reported in Tables 3 and 4, where the enrichment area width Δ has been varied in between 50 and 500 mm. For the sake of comparison it is recalled that the system condition number for the linear elastic case (no regularized discontinuity) is 2.0E + 6. Figure 9 reports the displacement contours at solution, that are undistinguishable for all the examined cases.

Table 3 reports the condition number in the case of absence of stabilization. With respect to the linear elastic case (no discontinuity) the condition number appears six orders of magnitude larger for $\Delta = 50$ and diverges to numerical infinity for $\Delta = 500$.

Table 3 Values of the system condition number for varying enrichment domain width in the 3D cut with regularized enrichment example and no stabilization

Δ	50	100	200	300	400	500
Condition number	8E + 12	7E + 13	1E + 15	8E + 16	4E + 19	∞

Table 4 Values of the system condition number for varying enrichment domain width in the 3D cut with regularized enrichment example and penalty stabilization

ω	1E + 3	1E + 4	1E + 5	1E + 6
$\Delta = 50$	2E + 7	2E + 7	2E + 7	5E + 7
$\Delta = 100$	2E + 7	2E + 7	2E + 7	5E + 7
$\Delta = 200$	2E + 7	2E + 7	2E + 7	5E + 7
$\Delta = 300$	2E + 7	2E + 7	2E + 7	5E + 7
$\Delta = 400$	2E + 7	2E + 7	2E + 7	5E + 7
$\Delta = 500$	2E + 7	2E + 7	2E + 7	5E + 7

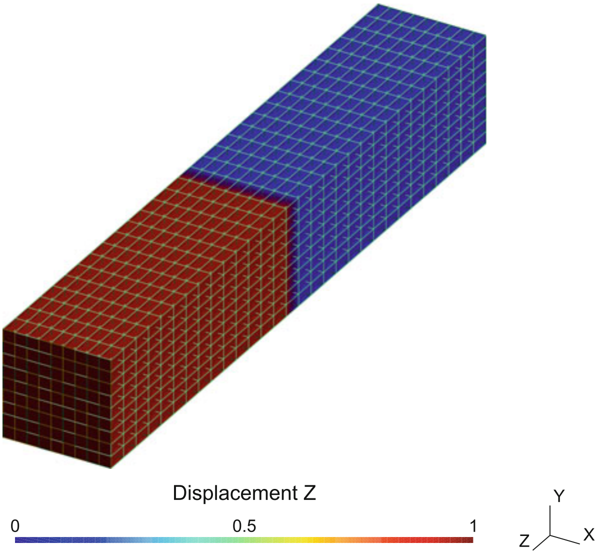


Fig. 9 Displacement contours for the regularized cut problem

On the other hand, when stabilization is introduced, Table 4, a very limited increase in the condition number is observed. The condition number remains practically constant for a wide range of the stabilization parameter scaling factor ω despite of the fact that the enrichment width is varied. This shows the high effectiveness and robustness of the proposed methodology. Assuming as optimal value for the stabilization parameter the largest possible value before affecting the condition number, from Table 4 the value $\omega = 1\text{E} + 5$ is derived, similarly to the previously analyzed case of pure jump and single element enrichment.

Similarly to the previous case of exact Heaviside enrichment, penalty stabilization introduces some low spurious stresses at solution. These are plotted in Fig. 10 for the limiting cases of regularisation lengths $\Delta = 50$ and $\Delta = 500$. Considering that when the discontinuity is absent (linear elastic case) the axial stress is equal to 8.3 MPa, from the contours in Fig. 10 an error smaller than 0.5 % is found, that is acceptable for technical applications. Stabilisation spurious stress can be reduced by increasing the stabilization parameter ω and accepting some impact on the system condition number.

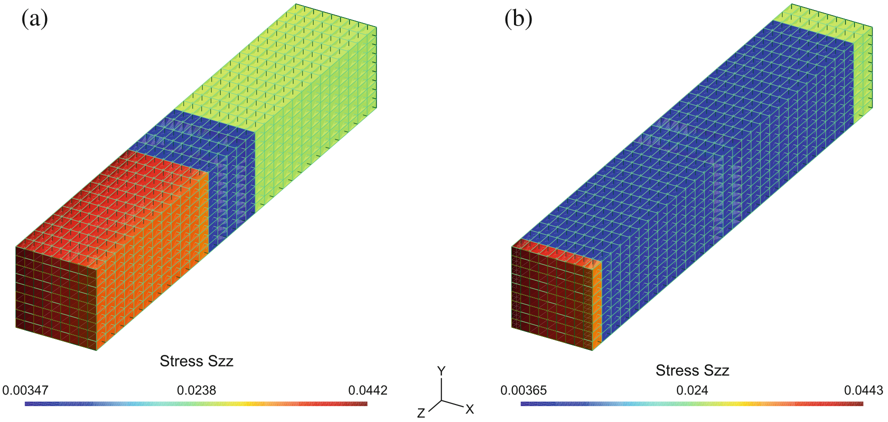


Fig. 10 Stabilisation spurious stresses at solution for $\Delta = 50$ (*left*) and $\Delta = 500$ (*right*). The *blue area* is associated to the enrichment width Δ

6 Conclusions

The addition of a stabilization penalty term biasing to zero the enrichment variables has proved to be an effective and very simple technique to eliminate indeterminacy in the system of equations of XFEM and XFEM-like formulations. In the paper the analytic structure of ill conditioning has been shown and numerical studies performed on 1D and 3D cases have evidenced a marked improvement in the condition number of the system of equations. Due to the extreme simplicity of the approach, the absence of additional computational cost and the possibility of leaving the basic problem formulation unchanged, the presented method appears a valid alternative to the existing techniques already present in the literature.

References

1. ABAQUS: ABAQUS Documentation. Dassault Systèmes, Providence (2014)
2. Abdelaziz, Y., Hamouine, A.: A survey of the extended finite element. Comput. Struct. **86**, 1141–1151 (2008)
3. Babuška, I., Banerjee, U.: Stable generalized finite element method (SGFEM). Comput. Methods Appl. Mech. Eng. **201–204**, 91–111 (2012)
4. Babuška, I., Banerjee, U., Osborn, J.: Survey of meshless and generalized finite element methods: a unified approach. Acta Numer. **12**, 1–125 (2003)
5. Béchet, E., Minnebo, H., Moës, N., Burgardt, B.: Improved implementation and robustness study of the X-FEM for stress analysis around cracks. Int. J. Numer. Methods Eng. **64**, 1033–1056 (2005)
6. Belytschko, T., Gracie, R., Ventura, G.: A review of extended/generalized finite element methods for material modeling. Model. Simul. Mater. Sci. Eng. **17**, p. Article number 043001 (2009)

7. Benvenuti, E.: Mesh–size–objective XFEM for regularized continuous–discontinuous transition. Finite Elem. Anal. Des. **47**, 1326–1336 (2011)
8. Benvenuti, E.: XFEM with equivalent eigenstrain for matrix–inclusion interfaces. Comput. Mech. **53**, 893–908 (2014)
9. Benvenuti, E., Tralli, A.: Simulation of finite-width process zone for concrete-like materials. Comput. Mech. **50**, 479–497 (2012)
10. Benvenuti, E., Tralli, A., Ventura, G.: A regularized XFEM model for the transition from continuous to discontinuous displacements. Int. J. Numer. Methods Eng. **74**, 911–944 (2008)
11. Benvenuti, E., Ventura, G., Ponara, N., Tralli, A.: Variationally consistent extended FE model for 3D planar and curved imperfect interfaces. Comput. Methods Appl. Math. Eng. **267**, 1–22 (2013)
12. Berger-Vergiat, L., Waisman, H., Hiriyur, B., Tuminaro, R., Keyes, D.: Inexact Schwarz-algebraic multigrid preconditioners for crack problems modeled by extended finite element methods. Int. J. Numer. Methods Eng. **90**, 311–328 (2012)
13. Duarte, C., Kim, D.: Analysis and applications of a generalized finite element method with global–local enrichment functions. Comput. Methods Appl. Mech. Eng. **197**, 487–504 (2008)
14. Ferté, G., Massin, P., Moës, N.: Convergence analysis of linear or quadratic X-FEM for curved free boundaries. Comput. Methods Appl. Mech. Eng. **278**, 794–827 (2014)
15. Fries, T.: A corrected XFEM approximation without problems in blending elements. Int. J. Numer. Methods Eng. **75**, 503–532 (2008)
16. Fries, T.P., Belytschko, T.: The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Methods Eng. **84**, 253–304 (2010)
17. Geuzaine, C., Remacle, J.-F.: Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. Int. J. Numer. Methods Eng. **79**, 1309–1331 (2009)
18. Gupta, V., Duarte, C.A., Babuška, I., Banerjee, U.: Stable GFEM (SGFEM): improved conditioning and accuracy of GFEM/XFEM for three-dimensional fracture mechanics. Comput. Methods Appl. Mech. Eng. **289**, 355–386 (2015)
19. Koehler, J.: On the dislocation theory of plastic deformation. Phys. Rev. **60**, 397–410 (1941)
20. Laborde, P., Pommier, J., Renard, Y., Salaun, M.: High order extended finite element method for cracked domains. Int. J. Numer. Meth. Eng. **64**, 354–381 (2005)
21. Lang, C., Makhija, D., Doostan, A., Maute, K.: A simple and efficient preconditioning scheme for heaviside enriched XFEM. Numer. Anal. arXiv:1312.6092v1 (2013)
22. Loehnert, S.: A stabilization technique for the regularization of nearly singular extended finite elements. Comput. Mech. **54**, 523–533 (2014)
23. LS-DYNA: User's Manual. Livermore Software Technology Corporation, Livermore (2013)
24. Martin, A., Esnault, J.-B., Massin, P.: About the use of standard integration schemes for X-FEM in solid mechanics plasticity. Comput. Methods Appl. Mech. Eng. **283**, 551–572 (2015):
25. Menk, A., Bordas, S.P.A.: A robust preconditioning technique for the extended finite element method. Int. J. Numer. Methods Eng. **85**, 1609–1632 (2011)
26. Mousavi, S.E., Pask, J.E., Sukumar, N.: Efficient adaptive integration of functions with sharp gradients and cusps in n-dimensional parallelepipeds. Int. J. Numer. Methods Eng. **91**, 343–357 (2012)
27. Müller, B., Kummer, F., Oberlack, M.: Highly accurate surface and volume integration on implicit domains by means of moment-fitting. Int. J. Numer. Methods Eng. **96**, 512–528 (2013)
28. Nabarro, F.: Theory of Crystal Dislocations. Dover Publications, New York (1967)
29. Sillem, A., Simone, A., Sluys, L.J.: The orthonormalized generalized finite element method-OGFEM: efficient and stable reduction of approximation errors through multiple orthonormalized enriched basis functions. Comput. Methods Appl. Mech. Eng. **287**, 112–149 (2015)
30. Strouboulis, T., Babuška, I., Copps, K.: The design and analysis of the generalized finite element method. Comput. Methods Appl. Mech. Eng. **181**, 43–69 (2000)
31. Sudhakar, Y., Wall, W.A.: Quadrature schemes for arbitrary convex/concave volumes and integration of weak form in enriched partition of unity methods. Comput. Methods Appl. Mech. Eng. **258**, 39–54 (2013)

32. Ventura, G.: On the elimination of quadrature subcells for discontinuous functions in the extended finite-element method. Int. J. Numer. Methods Eng. **66**, 761–795 (2006)
33. Ventura, G., Moran, B., Belytschko, T.: Dislocations by partition of unity. J. Numer. Methods Eng. **62**, 1463–1487 (2005)
34. Ventura, G., Gracie, R., Belytschko, T.: Fast integration and weight function blending in the extended finite element method. Int. J. Numer. Methods Eng. **77**, 1–29 (2009)
35. Ventura, G., Benvenuti, E.: Equivalent polynomials for quadrature in heaviside function enriched elements. Int. J. Numer. Methods Eng. **102**, 688–710 (2015)
36. Waisman, H., Berger-Vergiat, L.: An adaptive domain decomposition preconditioner for crack propagation problems modeled by XFEM. Int. J. Multiscale Comput. Eng. **11**, 633–654 (2013)
37. Yazid, A., Abdelkader, N., Abdelmadjid, H.: A state-of-the-art review of the X-FEM for computational fracture mechanics. Appl. Math. Model. **33**, 4269–4282 (2009)
38. Zhang, Q., Banerjee, U., Babuška, I.: Higher order stable generalized finite element method. Numer. Math. **128**, 1–29 (2014)

An Adaptive Fictitious Domain Method for Elliptic Problems

Stefano Berrone, Andrea Bonito, and Marco Verani

Abstract In the Fictitious Domain Method with Lagrange multiplier (`FDM`) the physical domain is embedded into a simpler but larger domain called the fictitious domain. The partial differential equation is extended to the fictitious domain using a Lagrange multiplier to enforce the prescribed boundary conditions on the physical domain while all the other data are extended to the fictitious domain. This lead to a saddle point system coupling the Lagrange multiplier and the extended solution of the original problem. At the discrete level, the Lagrange multiplier is approximated on subdivisions of the physical boundary while the extended solution is approximated on partitions of the fictitious domain. A significant advantage of the `FDM` is that no conformity between these two meshes is required. However, a restrictive compatibility conditions between the mesh-sizes must be enforced to ensure that the discrete saddle point system is well-posed. In this paper, we present an adaptive fictitious domain method (`AFDM`) for the solution of elliptic problems in two dimensions. The method hinges upon two modules `ELLIPTIC` and `ENRICH` which iteratively increase the resolutions of the approximation of the extended solution and the multiplier, respectively. The adaptive algorithm `AFDM` is convergent without any compatibility condition between the two discrete spaces. We provide numerical experiments illustrating the performances of the proposed algorithm.

S. Berrone
Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy
e-mail: tefano.berrone@polito.it

A. Bonito
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
e-mail: bonito@math.tamu.edu

M. Verani (✉)
MOX-Dipartimento di Matematica, Politecnico di Milano, P.zza L. da Vinci, I-20132 Milano, Italy
e-mail: marco.verani@polimi.it

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_11

1 Introduction

In many engineering applications the efficient numerical solution of partial differential equations on deformable or complex geometries is of paramount importance. In this respect, one crucial issue is the construction of the computational grid. To face this problem, one can basically resort to two different types of approaches. In the first approach, a mesh is constructed on a sufficiently accurate approximation of the exact physical domain (see, e.g., isoparametric finite elements [8], or isogeometric analysis [9]), while in the second approach one embeds the physical domain into a simpler computational mesh whose elements can intersect the boundary of the given domain. Clearly, the mesh generation process is extremely simplified in the second approach, while the imposition of boundary conditions requires extra work. Among the huge variety of methods sharing the philosophy of the second approach, let us mention here the Immersed Boundary methods (see, e.g., [16] and the references therein), the Penalty Methods (see, e.g., the seminal work [2]), the Fictitious Domain/Embedding Domain Methods (see, e.g., [7] and the references therein). In this paper, we focus on the Fictitious Domain Method with Lagrange multiplier (FDM) introduced in [11, 12] (see also [1] for the pioneering work inspiring this approach). In this approach, the physical domain ω is embedded into a simpler and larger domain Ω (the fictitious domain), the right-hand side is extended to the fictitious domain and the boundary conditions on the boundary of the physical domain are appended through the use of a Lagrange multiplier. The FDM gives rise to a saddle point problem whose exact solution restricted to ω corresponds to the solution of the original problem. At the discrete level, the FDM allows the use of structured and uniform meshes in the fictitious domain, without requiring any conformity between the bulk mesh and the boundary of the physical domain. This represents a relevant computational advantage. However, there are two important issues to be taken into account to build numerical techniques that are able to take advantage of the crucial features of FDM: the choice of the discrete spaces for the approximation of the solution and the multiplier, and the construction of the extension of the right-hand side from the physical domain to the fictitious one. As pointed out in the analysis performed in [11] in the context of finite element approximation for the solution of elliptic problems with Dirichlet boundary conditions, the first condition (*inf-sup* condition) is essential to ensure existence and uniqueness of the discrete solution, while the second one influences the regularity of the extended continuous solution, thus impacting on the approximation properties of the discrete spaces (on this latter topic, see, e.g., [14]). The first condition turns out to introduce some restrictive compatibility conditions between the mesh-sizes of the fictitious domain grid and the subdivision of boundary of the physical domain (needed to approximate the Lagrange multiplier). The second issue can spoil, for example, the performance of the linear finite element method on uniform meshes whenever the original solution is sufficiently regular, e.g. H^2 regular, while the extended solution is less regular.

In view of the above discussion, the computational effectivity of FDM seems to be a non trivial issue. However, as shown in the present paper, a judicious use of adaptivity can allow to overcome the above two obstructions and recover the full potentiality of FDM, i.e. working with discrete spaces (and meshes) violating the compatibility conditions and recover even in presence of less regular extended solutions, the optimal performance of finite elements on uniform grids in presence of regular solutions. In particular, in this work we present an adaptive fictitious domain method, named `AFDM`, based on the use of linear finite elements for the approximation of the solution and piecewise constants for the approximation of the Lagrange multiplier. In the spirit of the algorithm provided in [6], the method hinges upon two modules, `ELLIPTIC` and `ENRICH` that iteratively modify the discrete spaces for the approximation of the extended solution and the multiplier. Our method is proved to be convergent regardless of the imposition of any compatibility condition between the two discrete spaces. Similar remarks has been already pointed out in different contexts by Dahlke et al. [10] and Bänsch et al. [4] (see also [3] for an abstract discussion of inexact Uzawa methods) and we refer to [13] for the mathematical study of an adaptive algorithm for the Stokes system, which serves as a benchmark for saddle point problems. Moreover, preliminary numerical results show that `AFDM` is optimal with respect to the number of degrees of freedom employed to approximate the extended solution and the Lagrange multiplier. In two dimension, the optimality of the adaptive refinement strategy seems to require an adaptive strategy to generate the successive subdivision of the fictitious domain, while uniform or quasi-uniform subdivisions can be used for the boundary mesh.

The outline of the paper is as follows. In Sect. 2 we introduce the fictitious domain method, while in Sect. 3 we introduce the adaptive fictitious domain method. Finally, in Sect. 4 we numerically explore the convergence (and optimality) properties of our algorithm.

2 Fictitious Domains Method

Let ω be a bounded domain of $\mathbb{R}^2$ with boundary γ. To make the presentation simpler, we assume that ω is a polygon. We are interested in employing the fictitious domain method to solve the following model problem: let $f \in L^2(\omega)$, find $u \in H^1_0(\omega)$ such that

$$-\Delta u = f \qquad \text{in } \omega\ , \tag{1}$$

$$u = 0 \qquad \text{on } \gamma\ . \tag{2}$$

The fictitious domain formulation of problem (1)–(2) hinges on a square or rectangular domain Ω with boundary $\Gamma := \partial\Omega$ and such that $\omega \subset\subset \Omega$. It reads: for any L^2-extension $\tilde{f}$ of f to Ω, find $(\tilde{u}, \lambda) \in H_0^1(\Omega) \times H^{-\frac{1}{2}}(\gamma)$ such that

$$\int_\Omega \nabla\tilde{u} \cdot \nabla v - \langle \lambda, v \rangle_\gamma = \int_\Omega \tilde{f} v \, dx \qquad \forall v \in H_0^1(\Omega) \,, \tag{3}$$

$$\langle \mu, \tilde{u} \rangle_\gamma = 0 \qquad \forall \mu \in H^{-\frac{1}{2}}(\gamma) \,, \tag{4}$$

where for $v \in H_0^1(\Omega)$, its restriction to γ is understood in the sense of traces, and $\langle \cdot, \cdot \rangle_\gamma$ denotes the duality pairing between $H^{-\frac{1}{2}}(\gamma)$ and $H^{\frac{1}{2}}(\gamma)$ (recall that γ is a closed curve). Using an integration by parts formula, it is immediate to verify (see, e.g., [12]), that the fictitious domain formulation (3)–(4) is equivalent to original formulation (1)–(2) where

$$\lambda = \left[\frac{\partial\tilde{u}}{\partial n}\right]_\gamma \tag{5}$$

is the jump of $\frac{\partial\tilde{u}}{\partial n}$ across γ and n denotes the unit normal exterior to ω.

The regularity of $\tilde{u}$ and λ depends on the extension chosen for f and the domain ω. In the worst case, $\tilde{u} \in H^{\frac{3}{2}-\epsilon}(\Omega)$ for any $\epsilon > 0$ and $\lambda \in L^2(\gamma)$ satisfies $\lambda \in H^{\frac{1}{2}}(\gamma_i)$, for every straight line γ_i composing γ (see [11]). In what follows we will drop the symbol $\tilde{}$ if no confusion arises.

In [11], the Babuska-Brezzi's theory is used to guarantee that problem (3)–(4) is well posed. In particular, the bilinear form $(u, v) \mapsto \int_\Omega \nabla u \cdot \nabla v$ is coercive on the set $\{v \in H_0^1(\Omega) : \langle \mu, v \rangle_\gamma = 0 \ \forall \mu \in H^{-1/2}(\gamma)\}$ and that the following inf-sup condition holds: there exists a constant $\kappa > 0$ such that

$$\inf_{\mu \in H^{-\frac{1}{2}}(\gamma)} \sup_{v \in H_0^1(\Omega)} \frac{\langle v, \mu \rangle_\gamma}{\|\mu\|_{H^{-\frac{1}{2}}(\gamma)} \|v\|_{H^1(\Omega)}} \geq \kappa \,. \tag{6}$$

3 Adaptive Fictitious Domain Method

In this section we present our adaptive fictitious domain method (`AFDM`) based on Uzawa iterations. In Sect. 3.1, we describe the infinite dimensional version of the algorithm, whereas in Sect. 3.2 we introduce its adaptive finite dimensional counterpart.

3.1 Infinite Dimensional Fictitious Domain Algorithm

We start this section by describing the infinite dimensional fictitious domain algorithm for solving (3)–(4). It consists of Uzawa-type successive iterations: Given $\alpha > 0$ and $\lambda_0 \in H^{-\frac{1}{2}}(\gamma)$ we seek, for $j \geq 1$,

$$u_j \in H_0^1(\Omega): \qquad \int_\Omega \nabla u_j \cdot \nabla v = \int_\Omega f v + \langle \lambda_{j-1}, v \rangle_\gamma \quad \forall v \in H_0^1(\Omega)\,, \tag{7}$$

$$\lambda_j \in H^{-\frac{1}{2}}(\gamma): \qquad (\lambda_j, \mu)_\gamma = (\lambda_{j-1}, \mu)_\gamma - \alpha\, (u_j, \mu)_\gamma \quad \forall \mu \in H^{-\frac{1}{2}}(\gamma) \tag{8}$$

where we denote by $(\cdot,\cdot)_\gamma$ the scalar product in $H^{-\frac{1}{2}}(\gamma)$ and where we used the identification of $L^2(\gamma)$ and have with a slight abuse of notation $H^{\frac{1}{2}}(\gamma) \subset H^{-\frac{1}{2}}(\gamma)$.

The Schur complement operator $S : H^{-\frac{1}{2}}(\gamma) \to H^{\frac{1}{2}}(\gamma) \subset H^{-\frac{1}{2}}(\gamma)$ defined as

$$S\lambda = u_\lambda\,, \tag{9}$$

where $u_\lambda \in H_0^1(\Omega)$ is the solution to

$$\int_\Omega \nabla u_\lambda \cdot \nabla v = \langle \lambda, v \rangle_\gamma \quad \forall v \in H_0^1(\Omega)\,.$$

The operator S is symmetric and positive definite [12] and is instrumental in the analysis of the Uzawa iterations. In fact, (7)–(8) can be written using S as

$$\lambda_j = (I - \alpha S)\lambda_{j-1} + \alpha u_f \qquad \text{in } H^{-\frac{1}{2}}(\gamma), \tag{10}$$

where $u_f \in H_0^1(\Omega)$ is given by

$$\int_\Omega \nabla u_f \cdot \nabla v = \int_\Omega f\, v \qquad \forall v \in H_0^1(\Omega).$$

It is immediate to verify that if $0 < \alpha < 2/\|S\|_{\mathscr{L}(H^{-\frac{1}{2}}(\gamma), H^{-\frac{1}{2}}(\gamma))}$, then

$$\beta := \|I - \alpha S\|_{\mathscr{L}(H^{-\frac{1}{2}}, H^{-\frac{1}{2}})} < 1 \tag{11}$$

and thus the infinite dimensional fictitious domain algorithm is convergent.

3.2 Adaptive Finite Dimensional Fictitious Domain Method

We now introduce our adaptive finite dimensional fictitious domain algorithm (**`AFDM`**) which iteratively builds a sequence of nested finite dimensional spaces to achieve a reduction of the approximation error between each iterative step. We start with an initial conforming subdivision $\mathscr{T}_0$ of Ω made of triangles and an initial subdivision of γ (made of segments). We assume that

$$\text{for each } T \in \mathscr{T}_0,\ \overset{\circ}{T} \cap \gamma \text{ is connected,} \tag{12}$$

which is automatically satisfied upon assuming that the initial subdivision $\mathscr{T}_0$ is sufficiently fine to capture the interface and could be enforced via uniform refinements without affecting the asymptotic performances of the algorithm. From now on, $j \geq 0$ will always denote the **`AFDM`** iteration counter. We denote by $\mathscr{T}_j$ and $\mathscr{S}_j$ the j-th conforming partitions of Ω and γ made of triangles and segments, respectively. The diameters of the elements $T \in \mathscr{T}_j$ and $\ell \in \mathscr{S}_j$ are denoted $h_T := \operatorname{diam}(T)$ and $h_\ell := \operatorname{diam}(\ell)$, respectively. We emphasize that the two partitions built by the adaptive algorithm described below are mutually independent and in particular no compatibility conditions between the two partitions is required. This is a crucial difference from the results in [11] (see Remark 2 below). The shape regularity constant of a generic subdivision $\mathscr{T}$ is $\max_{T\in\mathscr{T}} \frac{h_T}{\rho_T}$, where ρ_T is the diameter of the largest ball inside T. The shape regularity constant of a sequence of subdivision $\{\mathscr{T}_i\}_{i\geq 0}$ is

$$\sup_{i\geq 0} \max_{T\in\mathscr{T}_i} \frac{h_T}{\rho_T}.$$

Associated with conforming partitions $\mathscr{T}$ of Ω and $\mathscr{S}$ of γ, we introduce the finite dimensional spaces

$$\mathbb{V}_{\mathscr{T}} := \left\{v \in C^0(\overline{\Omega}) : v|_T \in \mathscr{P}^1(T)\ \forall T \in \mathscr{T}_j\right\} \cap H^1_0(\Omega)$$

and

$$\mathbb{M}_{\mathscr{S}} := \{w \in L^2(\gamma) : w|_\ell \in \mathscr{P}^0(\ell)\ \forall \ell \in \mathscr{S}\},$$

where for $k \in \mathbb{N}$, $\mathscr{P}^k(D)$ is the set of polynomials of degree k in D. In the following, we set $\mathbb{V}_j := \mathbb{V}_{\mathscr{T}_j}$ and $\mathbb{M}_j := \mathbb{M}_{\mathscr{S}_j}$.

The finite dimensional adaptive fictitious domain algorithm relies on two subroutines described now.

3.2.1 The Module **`ELLIPTIC`**

The module **`ELLIPTIC`** adaptively constructs approximations U_j of the exact solution u_j to (7). To describe this procedure, we let $u_j \in H^1_0(\Omega)$ satisfying

$$\int_\Omega \nabla u_j \cdot \nabla v = \int_\Omega f\, v + \int_\gamma \Lambda_{j-1}\, v \qquad \forall v \in H^1_0(\Omega)\,, \tag{13}$$

for a given $\Lambda_{j-1} \in \mathbb{M}_{j-1}$.

In contrast to (7) where $\lambda_{j-1} \in H^{-\frac{1}{2}}(\gamma)$, we note that Λ_{j-1} belongs to a finite dimensional subspace of $L^2(\gamma)$. Moreover, we observe that (13) is a weak formulation of

$$\begin{aligned} &-\Delta u_j = f \ \text{ in } \omega, \quad -\Delta u_j = f \quad \text{in } \Omega \setminus \overline{\omega}, \\ &[u_j] = 0 \quad \text{on } \gamma, \quad \left[\frac{\partial u_j}{\partial n}\right]_\gamma = \Lambda_{j-1} \quad \text{on } \gamma, \\ &u_j = 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{14}$$

If ε_j stands for an adjustable error tolerance, then the module **ELLIPTIC**,

$$(\mathscr{T}_j, U_j) = \mathtt{ELLIPTIC}(\mathscr{T}_{j-1}, \Lambda_{j-1}, \varepsilon_j),$$

constructs adaptively a refined mesh $\mathscr{T}_j$ of $\mathscr{T}_{j-1}$ such that the solution of the discrete elliptic problem

$$U_j \in \mathbb{V}_j : \qquad \int_\Omega \nabla U_j \cdot \nabla V = \int_\Omega f\, V + \int_\gamma \Lambda_{j-1}\, V \qquad \forall V \in \mathbb{V}_j\,, \tag{15}$$

approximate u_j within in the prescribed tolerance ε_j, i.e.

$$\|\nabla(u_j - U_j)\|_{L^2(\Omega)} \le \varepsilon_j\,. \tag{16}$$

The adaptive **ELLIPTIC** module iterates a classical strategy of the type

$$\mathtt{SOLVE} - \mathtt{ESTIMATE} - \mathtt{MARK} - \mathtt{REFINE}$$

until condition (16) is satisfied (see, e.g., [15]).

The following upper bound for the error of any inner-iterate is instrumental in `ESTIMATE`. A corresponding lower bound is also valid and we refer to [5] for their proofs.

Proposition 3.1 (Upper Bound for ELLIPTIC) *Let $u_j \in H_0^1(\Omega)$ be the solution to (13) and $U_j^i \in \mathbb{V}_i$ be the discrete Galerkin solution to (15) associated with any refinement $\mathscr{T}^i$ of $\mathscr{T}_{j-1}$. Assume that the initial subdivision satisfy (12). Then, there exists a positive constant K^* only depending on $\mathscr{T}^i$ through its shape-regularity constant such that the following a posteriori error estimate holds*

$$\|\nabla(u_j - U_j^i)\|^2_{L^2(\Omega)} \le K^* \sum_{T \in \mathscr{T}^i} \eta^i(T)^2 \tag{17}$$

where

$$\eta^i(T)^2 := h_T^2 \|R_T^i\|_{L^2(T)}^2 + h_T \sum_{e\ edge\ of\ T} \|J_e^i\|_{L^2(e)}^2 + h_T \|\Lambda_{j-1}\|_{L^2(\mathring{T}\cap\gamma)}^2 ,$$

with

$$R_T^i := (f + \Delta U_j^i)|_T, \qquad J_e^i := \begin{cases} [\frac{\partial U_j^i}{\partial n}]_e - \Lambda_{j-1} & on\ e \cap \gamma, \\ [\frac{\partial U_j^i}{\partial n}]_e & on\ e \setminus \gamma. \end{cases} \tag{18}$$

Here $[.]_e$ *is the jump across the edge e and* $\mathring{T}$ *is the topological interior of T.*

Remark 1 (Elliptic Estimator) We note that the non-standard terms containing Λ_{j-1} in the estimator $\eta^i(T)$ above both measure the discrepancy from the exact relation $\lambda = [\frac{\partial u}{\partial n}]_\gamma$, see (5). In particular, within an element T, ∇U_j^i is continuous so that

$$\|\Lambda_{j-1}\|_{L^2(\mathring{T}\cap\gamma)} = \| \left[\frac{\partial U_j^i}{\partial m}\right]_{\mathring{T}\cap\gamma} - \Lambda_{j-1}\|_{L^2(\mathring{T}\cap\gamma)},$$

where $[\frac{\partial U_j^i}{\partial m}]_{\mathring{T}\cap\gamma}$ denotes the jump of $\frac{\partial U_j^i}{\partial m}$ across $\mathring{T} \cap \gamma$ and m denotes one of the normal to $\mathring{T} \cap \gamma$.

3.3 The Module `ENRICH`

Let $\mathbb{V}_j^\gamma$ be the restriction of functions in $\mathbb{V}_j$. We denote by $\Pi_j : \mathbb{V}_j^\gamma \to \mathbb{M}_j$ the orthogonal L^2-projection onto $\mathbb{M}_j$. The module

$$\mathscr{S}_j = \texttt{ENRICH}(\mathscr{S}_{j-1}, U_j, \varepsilon_j) , \tag{19}$$

constructs a refinement $\mathscr{S}_j$ of $\mathscr{S}_{j-1}$ such that

$$\|(I - \Pi_j)U_j\|_{H^{-\frac{1}{2}}(\gamma)} \le C\varepsilon_j \tag{20}$$

for a given C independent of j.

The quantity $\|(I - \Pi_j)U_j\|_{H^{-\frac{1}{2}}(\gamma)}$ is not computed exactly but estimated as follows. Standard interpolation error estimates together with a trace estimate and the stability of the L^2-projection yield

$$\|(I - \Pi_j)U_j\|_{H^{-1/2}(\gamma)} = \sup_{\phi \in H^{1/2}(\gamma)} \frac{\int_\gamma (I - \Pi_j)U_j \phi}{\|\phi\|_{H^{1/2}(\gamma)}} = \sup_{\phi \in H^{1/2}(\gamma)} \frac{\int_\gamma (I - \Pi_j)U_j (\phi - \Phi)}{\|\phi\|_{H^{1/2}(\gamma)}}$$

$$
\begin{aligned}
&\le \sup_{\phi\in H^{1/2}(\gamma)} \frac{\|h_\ell^{1/2}(I-\Pi_j)U_j\|_{L^2(\gamma)}\|h_\ell^{-1/2}(\phi-\Phi)\|_{L^2(\gamma)}}{\|\phi\|_{H^{1/2}(\gamma)}}\\
&\le C\|h_\ell^{1/2}(I-\Pi_j)U_j\|_{L^2(\gamma)} \le C(\max_{\ell\in\mathscr{S}_j} h_\ell)^{1/2}\|(I-\Pi_j)U_j\|_{L^2(\gamma)}\\
&\le C(\max_{\ell\in\mathscr{S}_j} h_\ell)\|U_j\|_{H^{1/2}(\gamma)} \le C(\max_{\ell} h_\ell)\|U_j\|_{H^1(\Omega)},
\end{aligned}
$$

where $\Phi \in \mathbb{M}_j$ and C is a generic constant independent of j. Now the stability estimate $\|U_j\|_{H^1(\Omega)} \le C\|f\|_{L^2(\Omega)}$ imply

$$\|(I-\Pi_j)U_j\|_{H^{-1/2}(\gamma)} \le C(\max_{\ell} h_\ell)\|f\|_{L^2(\Omega)}. \tag{21}$$

Based on this, the **ENRICH** routine refines recursively all the elements of $\mathscr{S}_{j-1}$ until they all have a mesh-size smaller than $\frac{\varepsilon_j}{\|f\|_{L^2(\Omega)}}$. The resulting mesh is output of **ENRICH** since it satisfies (20).

3.4 The Module UPDATE

The discrete Lagrange multiplier is updated by the module **UPDATE**

$$\Lambda_j = \mathtt{UPDATE}(\mathscr{T}_j, \mathscr{S}_j, \Lambda_{j-1}, U_j, \alpha)\ , \tag{22}$$

which computes according to (8)

$$\Lambda_j \in \mathbb{M}_j : \qquad \Lambda_j = \Lambda_{j-1} - \alpha \Pi_j U_j\ . \tag{23}$$

We note that as $U_j \in \mathbb{V}_j$ its restriction to γ is an element of $\mathbb{V}_j^\gamma$ (by construction).

3.5 The AFDM Algorithm

We are now in a position to detail the iterative structure of **AFDM**. Each iteration of the algorithm consists of an inner solver employing **ELLIPTIC** in place of (7), followed by an application of the module **ENRICH** and by an update of the multiplier performed by the module **UPDATE**.

```
ADAPTIVE FICTITIOUS DOMAIN METHOD (AFDM)
Given the initial triangulations $\mathscr{T}_0$ and $\mathscr{S}_0$, and
the parameters $\alpha$, $\varepsilon_0 > 0$, $0 < \zeta < 1$ set $j = 1$ and
iterate:
1. Select any function $\Lambda_0 \in \mathbb{M}_0$.
2. Update $\varepsilon_j \leftarrow \zeta\varepsilon_{j-1}$.
3. Compute $(\mathscr{T}_j, U_j) = \mathbf{ELLIPTIC}(\mathscr{T}_{j-1}, \Lambda_{j-1}, \varepsilon_j)$.
4. Enrich $\mathscr{S}_j = \mathbf{ENRICH}(\mathscr{S}_{j-1}, U_j, \varepsilon_j)$.
5. Compute $\Lambda_j = \mathbf{UPDATE}(\mathscr{T}_j, \mathscr{S}_j, \Lambda_{j-1}, U_j, \alpha)$.
6. Update $j \leftarrow j + 1$.
7. Go to step (2).
```

The algorithm **AFDM** is convergent [5] as reported in the theorem below.

Theorem 1 (Convergence) *Let $\alpha > 0$ be such that (11) holds and assume that the initial subdivision satisfy (12). Let $(\mathscr{T}_j, \mathscr{S}_j, U_j, \Lambda_j)$ be the sequence of meshes and subdivision produced by **AFDM**. There exist positive constants C_1 and $\delta < 1$ depending on the shape regularity constant of $\{\mathscr{T}_j\}_{j \geq 0}$ such that*

$$\|\nabla(u - U_j)\|_{L^2(\Omega)} + \|\lambda - \Lambda_j\|_{H^{-\frac{1}{2}}(\gamma)} \leq C_1 \delta^j \, . \tag{24}$$

Remark 2 (Compatibility Condition) In [11] the authors obtain a priori error estimates for the finite element approximation of (3)–(4) under the assumption that the mesh size of the bulk triangulation is sufficiently large compared to the mesh size of the boundary triangulation, i.e. there holds a compatibility condition between the discrete spaces. This latter is a crucial requirement to prove the validity of a discrete inf-sup condition and thus the existence of the discrete solution. We emphasize that the convergence of our adaptive algorithm holds without enforcing any compatibility condition, thus possibly violating, the discrete inf-sup condition (see [4, 10] for similar results in different contexts).

The algorithm **AFDM** described above never ends until a stopping criterion is appended. In principle, we can resort to (24) to obtain an a priori estimate for the necessary number of iterations to reach a prescribed tolerance. However, this strategy is not implementable as the constant δ appearing in (24) is not accessible in practice. For this reason, we now provide an a-posteriori error estimate for the quantity $\|\nabla(u - U_j)\|_{L^2(\Omega)} + \|\lambda - \Lambda_{j-1}\|_{H^{-\frac{1}{2}}(\gamma)}$ which can be employed to stop **AFDM** while guarantying a prescribed approximation error.

Proposition 3.2 (Upper Bound for AFDM) *Let $(u, \lambda) \in H^1_0(\Omega) \times H^{-\frac{1}{2}}(\gamma)$ be the solution to (3)–(4) and $\{(U_j, \Lambda_{j-1})\}$ be the sequence of approximations produced by **AFDM**. Assume that the initial subdivision satisfy (12). Then there exists a constant*

C_2 depending on the shape regularity constant of $\mathscr{T}_j$ such that

$$\|\nabla(u - U_j)\|_{L^2(\Omega)} + \|\lambda - \Lambda_{j-1}\|_{H^{-\frac{1}{2}}(\gamma)} \le C_2 \left(\eta_{\mathscr{T}_j} + \eta_{\mathscr{S}_{j-1}}\right) , \tag{25}$$

where

$$\eta_{\mathscr{T}_j} := \Big(\sum_{T\in\mathscr{T}_j} \eta_j(T)^2 \Big)^{1/2} \tag{26}$$

$$\eta_{\mathscr{S}_{j-1}} := \Big(\sum_{\ell\in\mathscr{S}_{j-1}} h_\ell \|\nabla_\gamma U_j\|^2_{L^2(\ell)} \Big)^{1/2} \tag{27}$$

and

$$\eta_j(T)^2 := h_T^2 \|R_{T,j}\|^2_{L^2(T)} + h_T \sum_{e \text{ edge in } T} \|J_{e,j}\|^2_{L^2(e)} + h_T \|\Lambda_{j-1}\|^2_{L^2(\mathring{T}\cap\gamma)},$$

with

$$R_{T,j} := (f + \Delta U_j)|_T, \qquad J_{e,j} := \begin{cases} [\frac{\partial U_j}{\partial n}]_e - \Lambda_{j-1} & \text{on } e \cap \gamma, \\ [\frac{\partial U_j}{\partial n}]_e & \text{on } e \setminus \gamma. \end{cases} \tag{28}$$

Remark 3 (Boundary Indicator) In view of the results contained in [17], the indicator $\eta_{\mathscr{S}_{j-1}}$ measures the mismatch between the trace of U_j on γ and the exact homogeneous Dirichlet boundary condition (2).

4 Numerical Results

In this section we illustrate numerically the convergence properties of the **`AFDM`** algorithm and investigate numerically its optimality studied in [5]. Before presenting the numerical results we discuss some details regarding the implementation of **`AFDM`**.

4.1 Implementation Issues

Data structures for representing the (two dimensional) triangular bulk mesh of the fictitious domain Ω and the (one dimensional) boundary mesh of γ are organized as binary trees starting from the 0-th level meshes $\mathscr{T}_0$ and $\mathscr{S}_0$, respectively. The initial bulk mesh $\mathscr{T}_0$ is a regular mesh of the domain Ω constructed ignoring the conformity with γ. The initial boundary mesh $\mathscr{S}_0$ is made of the edges of γ,

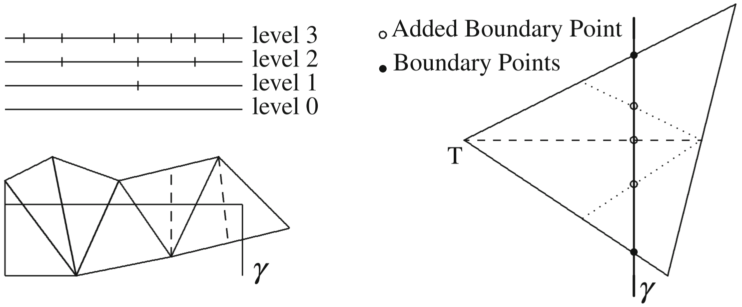


Fig. 1 Example of boundary refinement: (*top-left*) four different levels of refinement of the boundary mesh and (*bottom-left*) associated bulk mesh with local refinement (*dotted line*); (*right*) boundary points produced by the refinement of the bulk elements

which is assumed to be a polygonal curve, see Fig. 1(left). The refinement of the triangular bulk mesh is performed by employing the longest edge splitting. The refined elements are labeled as non-active and the newly created elements become active. Notice that additional refinements of neighboring elements may be needed in order to ensure conformity.

The refinement strategy advocated for the boundary mesh is more involved. On each boundary edge, we consider the points given by the intersection between the edge itself and the triangular mesh elements. We refer to these points as to boundary points. Whenever a refinement of a boundary edge is needed, we split the edge at the closest boundary point to the edge mid-point, see Fig. 1(right). In Fig. 1(bottom-left) we sketch three successive refinement of the horizontal portion of γ. In Fig. 1(top-left), we report the meshes resulting from the first four steps of a uniform refinement of the boundary elements. In case no boundary nodes are available, we perform successive refinements of the triangle containing the edge until one boundary nodes becomes available. However the two children produced by refinement of the triangle are labeled as non-active unless needed by the bulk mesh refinement procedure.

In our implementation the elements of the mesh $\mathscr{S}$ are always determined by the intersection between the boundary γ and a triangle (not necessarily active) belonging to the infinite binary tree with root $\mathscr{T}_0$. In view of this construction, the levels of refinement to which the elements in $\mathscr{S}$ and in $\mathscr{T}$ belong to are completely independent. Clearly, the above construction of $\mathscr{S}$ adds some geometric restrictions on the set of meshes we can deal with. However, the advantage at the computational level is quite remarkable because all the computer operations required by the refining process, as well as all the numerical computation of integrals involving the interaction between bulk and boundary objects (mesh elements or functions) can be always performed at the level of the single element (or of the binary tree stemming from it) *without* affecting the binary data structure of the neighboring elements.

For the sake of presentation in the rest of this section we employ the following notation. We set $e_j^u = u - U_j$ and $e_j^\lambda = \lambda - \Lambda_j$. Moreover, we denote by $\|\cdot\|_0$ the $L^2(\Omega)$ or $L^2(\gamma)$-norm depending on the context and by $|\cdot|_1$ the $H^1(\Omega)$-seminorm.

4.2 Test Case: L-Shape Domain

We consider the L-shaped domain $\omega = (-1\ 1)^2 \setminus (-1\ 0)^2$ with boundary γ. We are interested in the following model problem: we choose $f \in L^2(\omega)$ such that, after introducing polar coordinates (r, ϕ), the solution to (1)–(2) is

$$u(r, \phi) = h(r) r^{2/3} \sin(2/3(\phi + \pi/2))$$

where

$$h(r) = \frac{w(3/4 - r)}{w(r - 1/4) + w(3/4 - r)}, \quad w(r) = \begin{cases} r^2 & \text{if } r > 0 \\ 0 & \text{else.} \end{cases}$$

The fictitious domain formulation of problem (3)–(4) is obtained by embedding ω in the square domain $\Omega = (-1\ 1)^2$ with boundary $\Gamma = \partial\Omega$, see Fig. 2, and extend f by zero outside ω. It is not difficult to see that $u \in H^{\frac{5}{3}-\epsilon}(\omega)$, for any $\epsilon > 0$ and the exact Lagrange multiplier is

$$\lambda = \frac{2}{3} h(r)\, r^{-1/3}.$$

In the following, we explore the convergence and optimality properties of **AFDM** algorithm. The **AFDM** algorithm is applied with the following parameters: $\alpha = 0.5$, $\zeta = 0.95$ and different values of ε_0, namely $\varepsilon_0 = 1.0,\ 0.5,\ 0.25,\ 0.1$ (see Sect. 3 for the precise meaning of the parameters). In the sequel, we report the results obtained by **AFDM**. The outer iteration of **AFDM** is stopped when

$$\eta_{\mathcal{T}_j} + \eta_{\mathcal{S}_{j-1}} < \zeta^{45} \varepsilon_0.$$

In Fig. 2 we display the initial and the final mesh together with two intermediate adaptively refined meshes, while in Fig. 3 (left) we report the final discrete solution U plotted on the fictitious domain and in Fig. 3 (right) we collect the graphs of final discrete Lagrange multiplier Λ together with the exact multiplier λ and the restriction of U on the boundary γ. A close inspection of the figures reveals that the algorithm **AFDM** correctly approximates the exact pair (u, λ). As the value of β is unknown, an a priori choice for ξ is impossible. However, as already observed in [4], a practical choice for ξ seems to be $\xi = 0.95$. Indeed, larger values of ξ clearly guarantee the boundedness of the inner iterations of **ELLIPTIC**, but at the expense of an increased number of outer iterations. Values of ξ close to 0.95 ensure

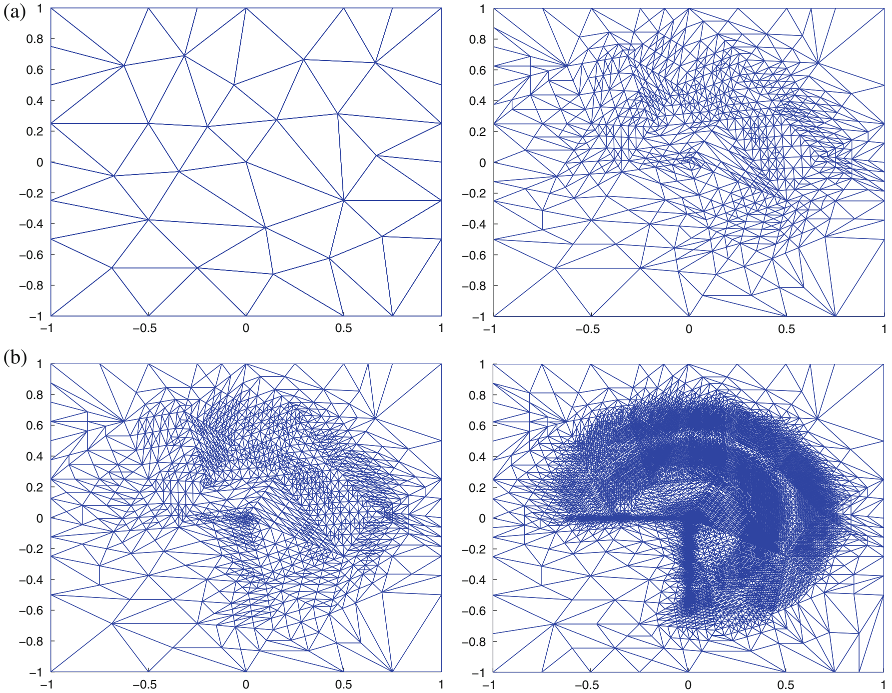


Fig. 2 Adaptive meshes produced by **`AFDM`**. (**a**) Initial mesh $\mathscr{T}_0$ (*left*) and mesh at iteration 12 (*right*), (**b**) mesh at iteration 18 (*left*) and final mesh (*right*)

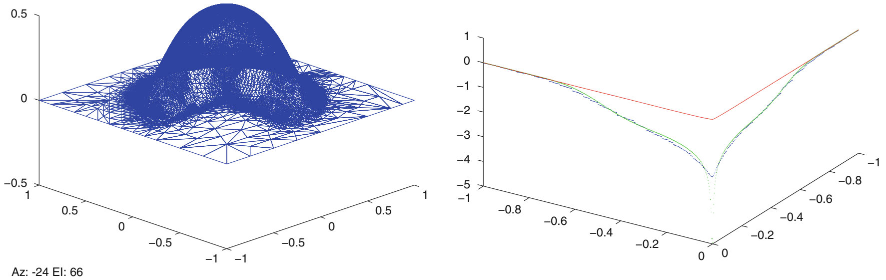


Fig. 3 *Left*: discrete solution U in the fictitious domain. *Right*: zoom on γ at the reentrant corner of the discrete solutions U (*red*), Λ (*blue*) and exact multiplier λ (*green*)

Table 1 Computed rates of convergence of the true errors and of the error estimators $\eta_{\mathscr{T}_j}$ and $\eta_{\mathscr{S}_{j-1}}$ with respect to the total number of dofs $\#\mathscr{T}_j + \#\mathscr{S}_{j-1}$ for different values of ε_0

	$\|e_j^u\|_0$	$\|e_j^u\|_1$	$\|e_{j-1}^\lambda\|_{-\frac{1}{2},j-1}$	$\eta_{\mathscr{T}_j}$	$\eta_{\mathscr{S}_{j-1}}$
$\varepsilon_0 = 1$	−0.9073	−0.5257	−0.4279	−0.5201	−0.6892
$\varepsilon_0 = 0.5$	−0.9347	−0.5447	−0.4232	−0.5080	−0.7072
$\varepsilon_0 = 0.25$	−0.9019	−0.5368	−0.4104	−0.5062	−0.7491
$\varepsilon_0 = 0.1$	−0.8819	−0.5584	−0.3924	−0.4980	−0.7966

Table 2 Growing rate r of the total number of dofs (second column), bulk dofs (third) and boundary dofs (fourth) with respect to ε_j, i.e. r such that $\#(\cdot)_j \in \mathscr{O}(\varepsilon_j^r)$

	$\#\mathscr{T}_j + \#\mathscr{S}_{j-1}$	$\#\mathscr{T}_j$	$\#\mathscr{S}_{j-1}$
$\varepsilon_0 = 1$	−1.8951	−1.9415	−0.9895
$\varepsilon_0 = 0.5$	−1.9326	−1.9595	−0.9928
$\varepsilon_0 = 0.25$	−2.0156	−2.0319	−1.0240
$\varepsilon_0 = 0.1$	−2.2928	−2.3096	−1.0831

an optimal decay rate of the H^1-error. This motivates the initial choice of setting $\xi = 0.95$ for running the numerical tests.

In the following, we further explore the optimality properties of **`AFDM`** and see their dependency on the parameter ε_0. In particular, in Table 1 we report the rates of convergence with respect to the total number of degrees of freedom, of the true errors for e_j^u in $L^2(\Omega)$, $H^1(\Omega)$, the error for e_{j-1}^λ in an approximate $H^{-\frac{1}{2}}(\gamma)$ computed as a weighted L^2 norm

$$\|e_{j-1}^\lambda\|_{-\frac{1}{2}} \approx \|e_{j-1}^\lambda\|_{-\frac{1}{2},j-1} := \Big(\sum_{\ell \in \mathscr{S}_{j-1}} h_\ell \|\lambda - \Lambda_{j-1}\|^2_{L^2(\ell)}\Big)^{1/2},$$

and the error estimators $\eta_{\mathscr{T}_j}$ and $\eta_{\mathscr{S}_{j-1}}$.

From the third and fourth columns we infer that the total error $\|\nabla(u - U_j)\|_{L^2(\Omega)} + \|\lambda - \Lambda_{j-1}\|_{H^{-\frac{1}{2}}(\gamma)}$ decreases to zero roughly as $(\#\mathscr{T}_j + \#\mathscr{S}_{j-1})^{-0.5}$, the optimal decay for piecewise approximations of u. This is corroborated by the results in columns five and six which are thus in agreement with Proposition 3.2.

In Table 2 we collect the growing rates of the total number of dofs $\#\mathscr{T}_j + \#\mathscr{S}_{j-1}$, of the bulk dofs $\#\mathscr{T}_j$ and of the boundary dofs $\#\mathscr{S}_{j-1}$. The results shows that optimal convergence is dictated by the regularity of the original non-extended solution u. Finally, collecting the results of Tables 1–2, we realize that both the total error and the error indicator $\eta_{\mathscr{T}_j} + \eta_{\mathscr{S}_{j-1}}$ decay as ε_j.

5 Conclusions

We introduced an Adaptive Fictitious Domain Method (**`AFDM`**). The core of the method is based on two modules, **`ELLIPTIC`** and **`ENRICH`** that iteratively modify the discrete spaces for the approximation of the extended solution and the multiplier, respectively. Numerical results show that **`AFDM`** is convergent, regardless of the imposition of any compatibility condition between the two discrete space. Moreover, preliminary tests seem to suggest the optimal behaviour of **`AFDM`**. This last topic requires further investigations, which is the topic of [5].

Acknowledgements This work has been partially supported by the Italian MIUR through PRIN research grant 2012HBLYE4_001 "Metodologie innovative nella modellistica differenziale numerica", by INdAM-GNCS and by the National Science Foundation grant DMS-1254618.

References

1. Babuška, I.: The finite element method with Lagrangian multipliers. Numer. Math. **20**, 179–192 (1972/1973)
2. Babuška, I.: The finite element method with penalty. Math. Comput. **27**, 221–228 (1973)
3. Bacuta, C.: A unified approach for Uzawa algorithms. SIAM J. Numer. Anal. **44**(6), 2633–2649 (2006)
4. Bänsch, E., Morin, P., Nochetto, R.H.: An adaptive Uzawa FEM for the Stokes problem: convergence without the inf-sup condition. SIAM J. Numer. Anal. **40**(4), 1207–1229 (2002)
5. Berrone, S., Bonito, A., Verani, M.: An adaptive fictitious domain method for elliptic problems: convergence and optimality (2016, in preparation)
6. Bonito, A., DeVore, R.A., Nochetto, R.H.: Adaptive finite element methods for elliptic problems with discontinuous coefficients. SIAM J. Numer. Anal. **51**(6), 3106–3134 (2013)
7. Börgers, C., Widlund, O.B.: On finite element domain imbedding methods. SIAM J. Numer. Anal. **27**(4), 963–978 (1990)
8. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Classics in Applied Mathematics, vol. 40. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2002). Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)]
9. Cottrell, J.A., Hughes, T.J.R., Bazilevs, Y.: Isogeometric Analysis: Toward Integration of CAD and FEA. Wiley, Hoboken (2009)
10. Dahlke, S., Dahmen, W., Urban, K.: Adaptive wavelet methods for saddle point problems—optimal convergence rates. SIAM J. Numer. Anal. **40**(4), 1230–1262 (2002)
11. Girault, V., Glowinski, R.: Error analysis of a fictitious domain method applied to a Dirichlet problem. Japan J. Indust. Appl. Math. **12**(3), 487–514 (1995)
12. Glowinski, R., Pan, T.-W., Périaux, J.: A fictitious domain method for Dirichlet problem and applications. Comput. Methods Appl. Mech. Eng. **111**(3–4), 283–303 (1994)
13. Kondratyuk, Y., Stevenson, R.: An optimal adaptive algorithm for the Stokes problem. SIAM J. Numer. Anal. **46**(2), 747–775 (2008)
14. Mommer, M.S.: A smoothness preserving fictitious domain method for elliptic boundary-value problems. IMA J. Numer. Anal. **26**(3), 503–524 (2006)
15. Nochetto, R.H., Veeser, A.: Primer of adaptive finite element methods. In: Multiscale and Adaptivity: Modeling, Numerics and Applications. Lecture Notes in Mathematics, vol. 2040, pp. 125–225. Springer, Heidelberg (2012)
16. Peskin, C.S.: The immersed boundary method. Acta Numer. **11**, 479–517 (2002)
17. Sacchi, R., Veeser, A.: Locally efficient and reliable a posteriori error estimators for Dirichlet problems. Math. Models Methods Appl. Sci. **16**(3), 319–346 (2006)

Higher-Order Accurate Integration for Cut Elements with Chen-Babuška Nodes

Thomas-Peter Fries

Abstract The higher-order accurate numerical integration of geometries that are implicitly defined by level-set functions is considered. A higher-order background mesh is employed providing an interpolation of the level-set function by Lagrangian shape functions. The integration may take place on the zero-level set or in the domains defined by the sign of the level-set function. This work is a follow-up of Fries and Omerović (Int J Numer Methods Eng, doi:10.1002/nme.5121). Herein, it is shown that special distributions of the element nodes, which are optimized for integration, yield significantly better results than equally-spaced nodes. Different error norms are proposed which allow to investigate the accuracy of general implicit geometries in two and three dimensions.

1 Introduction

The integration of implicit geometries is a standard task in many numerical methods such as fictitious domain methods [3, 8, 18] and extended finite element methods (XFEM) [2, 9, 19]. In the context of the level-set method [23, 27], the boundaries of the geometry or the interfaces inside are described by the zero-level set of a scalar function in the domain. The level-set values are typically given at the *nodes* of a background mesh which does not consider for the position of the interfaces or boundaries. Herein, we assume a higher-order background mesh composed by Lagrangian elements which imply interpolation functions for the level-set function inside the elements. The zero-level sets are typically curved and arbitrarily cut the background elements.

It is well-known that *equally spaced* nodes in the higher-order Lagrange elements of the background mesh feature problems for interpolations [24, 25, 29], in particular they suffer from the Runge phenomenon. In [4, 5], optimal interpolation points are provided for standard elements in one, two, and three dimensions. Herein, these special node distributions are investigated and it is confirmed that the resulting

---

T.-P. Fries (✉)
Institute of Structural Analysis, Graz University of Technology, Lessingstr. 25/II, 8010 Graz, Austria
e-mail: fries@tugraz.at

G. Ventura, E. Benvenuti (eds.), *Advances in Discretization Methods*,
SEMA SIMAI Springer Series 12, DOI 10.1007/978-3-319-41246-7_12

integration points are considerably more accurate than with equally spaced element nodes as used in [10].

Two standard integration tasks occur in the context of the integration of implicit geometries: The integration on the zero-level set or on the two sides of the zero-level set. For example, in fictitious domain methods an integration on the zero-level set is necessary to account for boundary conditions and an integration on one side of the zero-level set to integrate the governing equations of a model in weak form. In the XFEM, an integration on the zero-level set may be needed to account for interface conditions and on the two sides of the interface to evaluate volume terms in the weak form.

A unified treatment of these two integration tasks in two and three dimensions was proposed by the authors of this work in [10] for equally spaced element nodes. This work is a follow-up which (1) extends the algorithm to Lagrange elements with special node distributions [4, 5], (2) extends the iterative method to locate positions on the zero-level set, and (3) proposes new error norms suitable for general implicit geometries. In [10], results have only been presented based on circles in two dimensions and spheres in three dimensions, and some of the suggested error norms are not useful for general level-set functions. Furthermore, this work may also be seen as a summary of the elaborate description proposed by the authors in [10].

Approaches for the integration of elements with internal boundaries and interfaces may be distinguished based on the fact whether they rely on a decomposition of the cut elements into sub-elements or not. For the first class, the standard approach is to recursively decompose a cut element into polygonal sub-cells until the desired accuracy is reached [1, 7, 20]. This typically leads to a very large number of integration points in the context of higher-order approximations [7, 13, 15, 28, 33]. It was already noted in [6, 10, 14, 26] that a decomposition into sub-elements with curved, higher-order edges or faces is a strategy which consistently enables to obtain higher-order accurate integration rules with only a modest number of integration points. Of course, the effort for generating these integration points is larger than for the polygonal approaches with reduced accuracy. The other class of approaches is built by methods that do not decompose the cut elements, the interested reader is referred to [21, 22, 31, 32]. Compared to the element decomposition techniques, it seems that these approaches are often less intuitive and general. Their extendability to three dimensions, higher-order accuracy, and general integrands is not guaranteed.

The paper is organized as follows: The two main sections describe the integration in two and three dimensions, respectively. In each section, the meshing of the zero-level set by higher-order interface elements inside the cut reference background elements is described. This is called reconstruction and is sufficient for the integration on zero-level sets. For the integration on the two sides of the zero-level set, sub-elements are defined on the two sides which feature a higher-order side coinciding with the reconstructed interface element. The blending function method is used for this purpose [11, 12, 30]. The paper concludes in Sect. 4.

2 Integration in Two Dimensions

Let us assume a two-dimensional domain $\Omega \in \mathbb{R}^2$ which is discretized by Lagrange elements of order m_Ω. These elements build the background mesh and their edges do not align with potential interfaces and boundaries. The element nodes are either equally spaced in the corresponding reference elements or they feature a special distribution optimized for interpolation. In [4], such nodes are given by Chen and Babuška in one-dimensional elements and triangles, see Figs. 1 and 2a, respectively. For quadrilaterals, the points are easily determined by building tensor-products of the one-dimensional distributions, see Fig. 2b. We label elements with the special nodes "Chen-Babuška elements" in this work, referring to the authors of [4, 5]. The shape functions of the Lagrange elements with equally spaced nodes or Chen-Babuška nodes share some important properties: (1) they feature Kronecker-δ property, (2) they are C_0-continuous across element boundaries, (3) they are negative in some parts of the element (yet bounded for Chen-Babuška elements but unbounded for equally spaced nodes). It is important to note that the Runge phenomenon is a well-known problem of Lagrange elements with equally spaced nodes but is largely reduced for Chen-Babuška elements.

There exists a continuous scalar function $\phi(\mathbf{x})$ for all $\mathbf{x} \in \Omega$, called the level-set function. The zero-level set is defined as

$$\Gamma_0 = \{\mathbf{x} \in \Omega : \phi(\mathbf{x}) = 0\}, \tag{1}$$

which is a (curved) one-dimensional line in two dimensions. This line may be interpreted as an *interface* which cuts the domain into two non-overlapping regions Ω^+ and Ω^- based on the sign of $\phi(\mathbf{x})$; this is frequently the case in applications of the XFEM. It may also define the *boundary* of a domain, then Ω^+ may be

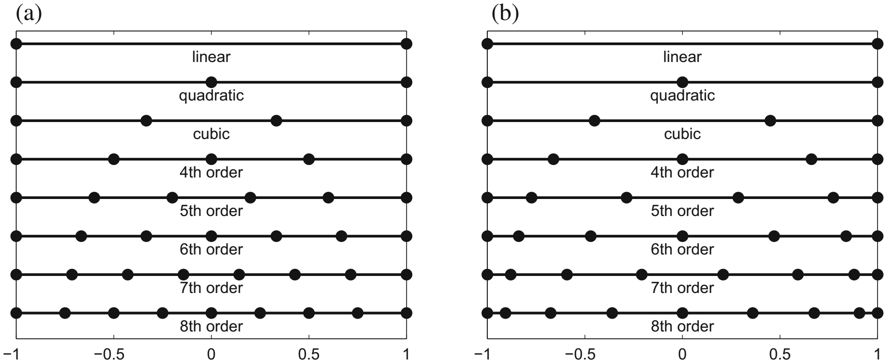


Fig. 1 Equally spaced nodes and Chen-Babuška nodes in one-dimensional intervals. (**a**) Equally spaced nodes. (**b**) Chen-Babuška nodes

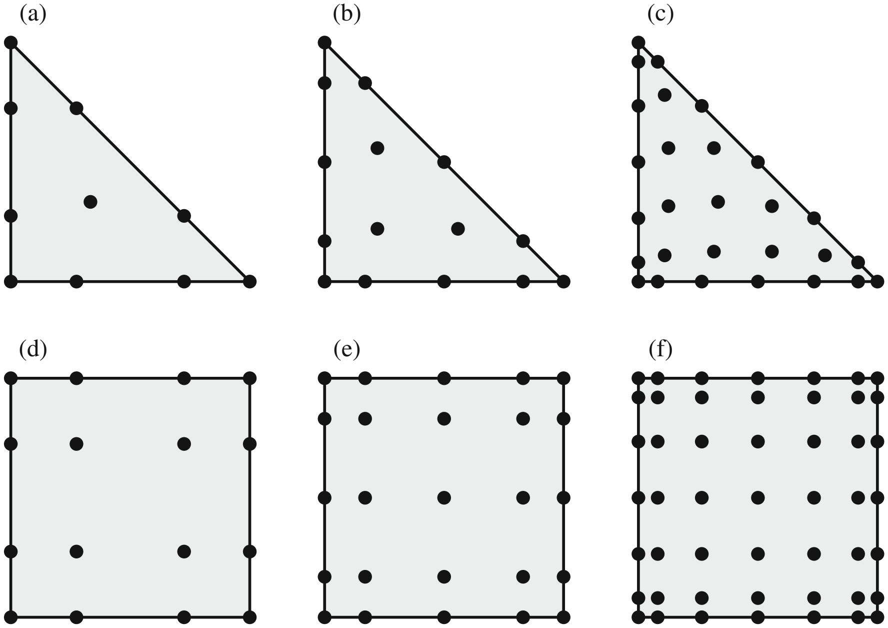


Fig. 2 Triangular and quadrilateral Chen-Babuška elements of different order. (**a**) Cubic, (**b**) 4th order, (**c**) 6th order, (**d**) cubic, (**e**) 4th order, (**f**) 6th order

the domain and Ω^- the outer region which is neglected in a simulation; this is frequently the case in applications of fictitious domain methods.

The level-set function $\phi(\mathbf{x})$ is replaced by its discrete counterpart

$$\phi^h(\mathbf{x}) = \sum_{i \in I} N_i^{\Omega}(\mathbf{x}) \cdot \phi_i \tag{2}$$

where I is the set of all nodes in Ω, $N_i^{\Omega}(\mathbf{x})$ are the higher-order shape functions, and $\phi_i = \phi(\mathbf{x}_i)$ are the nodal level-set values.

It is obvious that the zero-level set may cut arbitrarily through the elements of the background mesh. It is then a standard task, e.g. in applications of the XFEM and fictitious domain methods, to properly integrate on the two sides of the zero level-set or right on the zero-level set. The focus is now on *one* element cut by the zero-level set.

The procedure proposed in [10] is to first approximate the zero-level set of $\phi^h(\mathbf{x})$ by means of a higher-order, typically curved, one-dimensional interface element in the *reference* background element. This is called "reconstruction" and involves finding roots of the higher-order level-set function through an iterative procedure. It is then simple to integrate on the zero-level set in the domain Ω by first mapping the nodes of the curved interface element to the *physical* background element and then mapping one-dimensional Gauss points to this interface element, see Fig. 3.

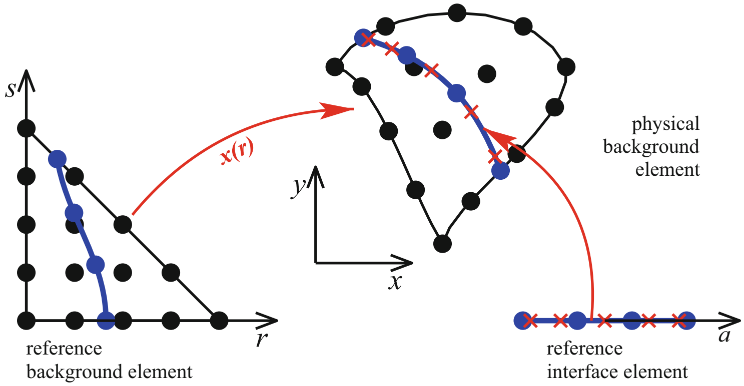


Fig. 3 Sketch of the integration on the zero-level set. The element nodes of the curved 1D interface element are located in the reference background element and mapped to the physical element. Then, standard Gauss points in 1D (*red crosses*) are mapped to this interface element

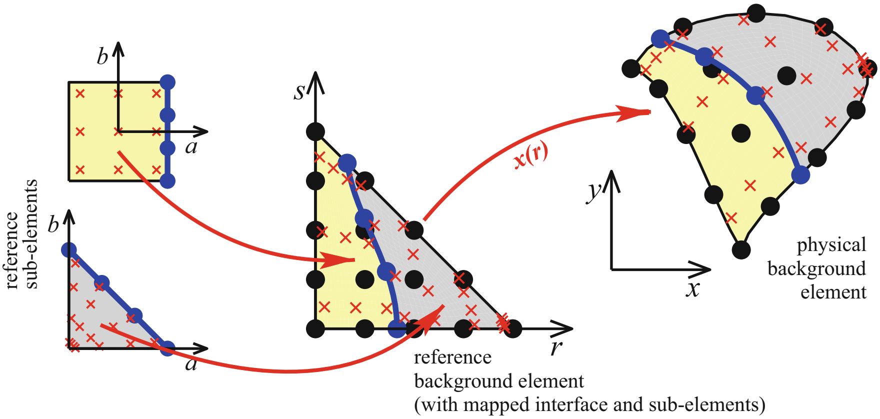


Fig. 4 This sketch demonstrates the integration in the cut elements. The reconstructed curved interface element cuts the reference background element into special sub-elements with one higher-order side. There, standard Gauss points (*red crosses*) are located and mapped to the reference background element and further to the physical element

For the integration on the two sides of the zero-level set, the background reference element is decomposed into sub-elements. Therefore, also the reconstructed interface element is needed and special sub-elements are defined which feature one higher-order side coinciding with the interface element. The blending function mapping is employed for this purpose [11, 12, 30]. Standard Gauss points in two dimensions are mapped to the special sub-elements in the reference background element and further to the physical background element, see Fig. 4.

It is obvious that higher-order elements may feature very complicated topologies of zero-level sets. For example, element edges may be cut several times, or the zero-

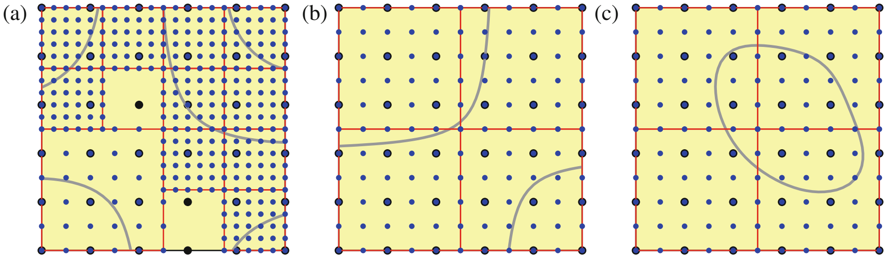


Fig. 5 Some complicated zero-level sets in quadrilateral elements corresponding to invalid level-set data. Note that after recursive refinements, valid level-set data is obtained in the refined elements. (**a**) Some edges are cut more than once. (**b**) More than two edges are cut. (**c**) The element is cut although none of the edges are cut

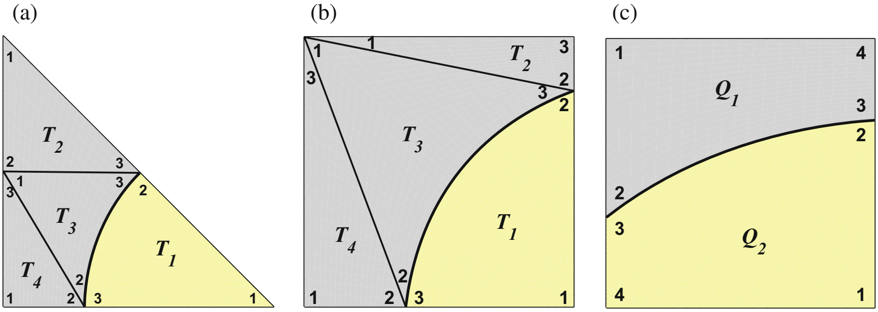


Fig. 6 Topologically different cut situations for valid level-set data in triangles and quadrilaterals. (**a**) Triangle, (**b**) quadrilateral, top. 1, (**c**) quadrilateral, top. 2

level set is completely inside an element, see Fig. 5. Therefore, a recursive procedure may be required (typically only in very few elements) until a "valid topology" inside the element is obtained. By "valid" we refer to the situation where (1) each element edge is only cut once, (2) the overall number of cut edges must be two, and (3) if no edge is cut then the element is completely uncut. Whether these conditions are fulfilled is checked based on a sample grid in the reference element. If they are not matched, the reference element is recursively refined and the described procedures apply to the refined elements. For an overview, see Fig. 5 and for details, see [10].

Provided that valid level-set data is obtained, the zero-level set cuts the reference element in a limited number of topologically different cases, see Fig. 6. Only the sign at the *corner* nodes of the reference element determines the situation. A cut triangle is decomposed into one sub-triangle and one sub-quadrilateral; the latter is further decomposed into three sub-triangles. For cut quadrilaterals, two different topologies are possible. Topology 1 is obtained when two neighboring edges are cut (resulting into four sub-triangles). Topology 2 is obtained when two opposite edges are cut leading to two sub-quadrilaterals. Note that the sub-elements feature one curved side which coincides with the reconstructed interface element.

2.1 Reconstruction of Zero-Isolines in 2D Background Elements

In order to approximate the zero-level set in the reference background element by (curved) interface elements of the desired order m_Γ, element nodes have to be located on the zero-level set. Therefore, an iterative procedure is required for each node. Many alternatives are possible, being different e.g. in the starting values and search paths. In [10], a Newton-Raphson based method was suggested using *straight* search paths. It was demonstrated that the issue of starting points and corresponding search paths has important consequences on the resulting accuracy of the integration schemes. Most importantly, starting values in two-dimensional elements should at least be provided by a linear reconstruction of the zero-level set. This is the straight line between the intersection points of the zero-level set and the element edges, see Fig. 7. For *inner* nodes of higher-order interface elements, good starting guesses are found by mapping the higher-order element nodes to the linear interface element. This works analogously for equally spaced nodes and Chen-Babuška nodes. In Fig. 7, the starting points for a Chen-Babuška element of order 5 are shown for different scenarios in triangles and quadrilaterals.

It is now necessary to address the question how to move towards suitable positions on the zero-level set from the starting points on the linear reconstruction. Two approaches relying on straight search paths are proposed in [10]: approach 1 uses the normal direction to the linear interface element (Fig. 8a), approach 2 uses the gradient of the level-set function $\phi(\mathbf{x})$ at the starting point (i.e. the path perpendicular to each contour line (Fig. 8b). Approach 3 is new and employs a search path which is not straight for the whole iterative procedure. It rather uses the gradient of $\phi(\mathbf{x})$ at each intermediate position during the iteration. All approaches yield quadratic convergence rates of the iterative procedure but, of course, to (slightly) different positions of the element nodes on the zero-level set.

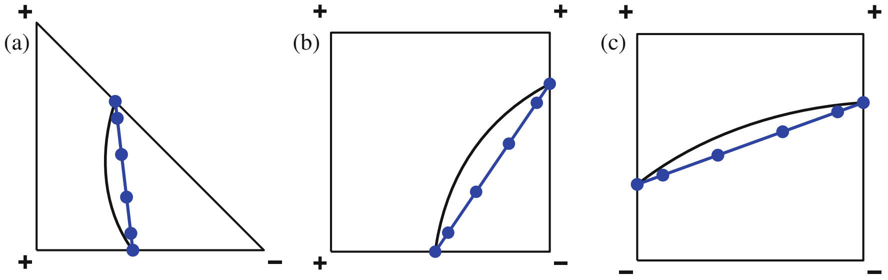


Fig. 7 Starting values for the root finding. The *thick black line* is the exact zero-level set of $\phi^h(\mathbf{x})$. The intersections with the element edges provide a straight line (linear reconstruction). Interface elements of the desired order m_Γ (with equally spaced nodes or Chen-Babuška nodes) are mapped to this linear element. (**a**) Triangle, (**b**) quadrilateral, top. 1, (**c**) quadrilateral, top. 2

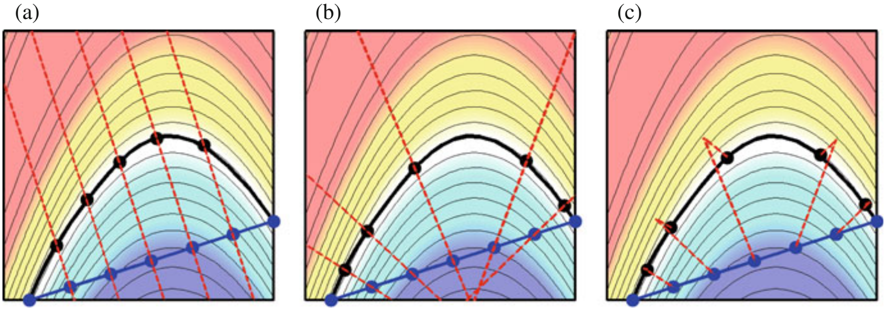


Fig. 8 Different approaches for search paths at the starting points on the linear reconstruction. (**a**) Approach 1, (**b**) approach 2, (**c**) approach 3

Mathematically, the algorithm is described as follows. The task is to find the root of $\phi^h(\mathbf{r})$ in the reference element, i.e. some position on the zero-level set. The starting point of the iterative procedure is labeled $\mathbf{r}^0$. The algorithm for all approaches considered here is based on the following iteration:

$$\mathbf{r}^{i+1} = \mathbf{r}^i - \frac{\phi^h\left(\mathbf{r}^i\right)}{\nabla\phi\left(\mathbf{r}^i\right) \cdot \mathbf{N}} \cdot \mathbf{N}. \tag{3}$$

In approach 1, **N** is the normal vector of the linear reconstruction which is equal for all starting points. In approach 2, **N** is the gradient of ϕ at each starting point, hence $\mathbf{N} = \nabla\phi\left(\mathbf{r}^0\right)$. Finally, approach 3 uses the current gradient of ϕ at $\mathbf{r}^i$, i.e. $\mathbf{N} = \nabla\phi\left(\mathbf{r}^i\right)$. When all nodes of the higher-order interface element are determined in the *reference* background element, it is a simple task to map them to the *physical* background elements using the isoparametric concept.

It is now clear how to reconstruct higher-order interface elements in the reference and physical background elements which approximate the zero-level set with the expected accuracy. The integration on the zero-level set is straightforward when mapping standard Gauss points in one-dimensional intervals to the curved interface elements, see Fig. 3.

2.2 *Integration in Cut 2D Background Elements*

For an integration in a cut element, the element is decomposed into special sub-elements which feature one higher-order side that coincides with the reconstructed interface element, see Fig. 4. A mapping of standard Gauss points in triangles and quadrilaterals to these special sub-elements has to be defined. The blending function mapping is used for this purpose; we refer to [10] for details and a unified assessment for the situation in different element types. The situation of Chen-Babuška elements is extended straightforwardly to the description therein. Also the issue of "consistent integration rules" in triangles applies just as well: It is crucial to use tensor-product Gauss rules of quadrilaterals also in the triangular sub-elements (collapsing at the node opposite to the higher-order side) in order to avoid a negative impact on the accuracy due to the properties of the blending function mapping in triangles. Namely, the Jacobian is not smooth at the node opposite to the higher-order side.

2.3 Numerical Results in Two Dimensions

Convergence studies are presented for the integration in two-dimensional domains: Firstly, on the zero-level set of $\phi(\mathbf{x})$ and, secondly, in the region where $\phi(\mathbf{x}) < 0$, i.e. Ω^-. The domain is discretized by higher-order background meshes with equally spaced nodes or Chen-Babuška nodes in each element, respectively. It is noted that also the reconstructed interface elements may feature (independently of the background elements) equally spaced nodes or Chen-Babuška nodes. Therefore, four different setups are distinguished in Table 1, see also Fig. 9.

It was shown in [10] that the accuracy of the integration is directly related to the order of the background *and* interface elements, so that it is natural to choose identical orders for both, i.e. $m_\Omega = m_\Gamma$. Following [10], triangular and quadrilateral elements are investigated here, each in a Cartesian arrangement where the element edges align with the physical coordinate system ("undeformed mesh") and in a deformed mesh where the elements are perturbed, see Fig. 10. Two different

Table 1 Combinations of equally spaced nodes and Chen-Babuška nodes in background and interface elements

	Background elements	Interface elements
Case 1: ES-ES	Equally spaced nodes	Equally spaced nodes
Case 2: ES-CB	Equally spaced nodes	Chen-Babuška nodes
Case 3: CB-ES	Chen-Babuška nodes	Equally spaced nodes
Case 4: CB-CB	Chen-Babuška nodes	Chen-Babuška nodes

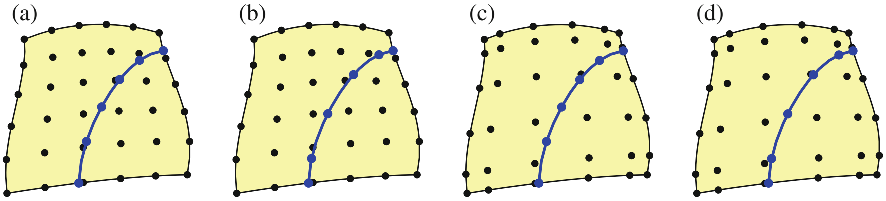


Fig. 9 Equally spaced (ES) and Chen-Babuška (CB) nodes in background elements and reconstructed interface elements (of order 5). (**a**) Case 1: ES-ES. (**b**) Case 2: ES-CB. (**c**) Case 3: CB-ES. (**d**) Case 4: CB-CB

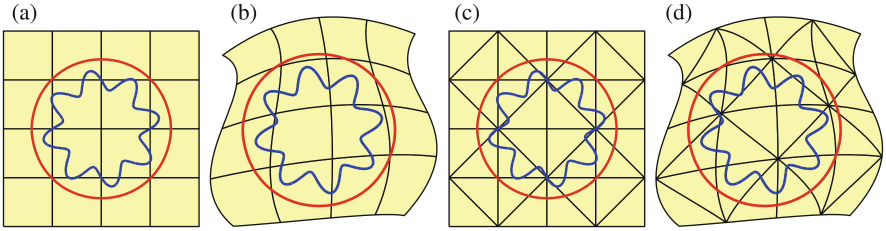


Fig. 10 Deformed and undeformed background meshes composed by higher-order triangular and quadrilateral elements. The *red* and *blue lines* are the exact zero-level sets of $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$, respectively. (**a**) Undef., quad., (**b**) def., quad., (**c**) undef., tri., (**d**) def., tri.

level-set functions are considered:

$$\phi_1(\mathbf{x}) = \sqrt{x^2 + y^2} - r \tag{4}$$

$$\phi_2(\mathbf{x}) = \sqrt{x^2 + y^2} - R(\theta) \tag{5}$$

with $r = 0.7123$, $R(\theta) = 0.5 + 0.1 \cdot \sin(8\theta)$ and $\theta(\mathbf{x}) = \operatorname{atan}(y/x)$. See Fig. 10 for a visualization of the zero-level sets of ϕ_1 and ϕ_2. The zero-level set of ϕ_1 is a circle with radius r and is frequently used in the literature, e.g. [10], and ϕ_2 is similar to [16].

In the convergence studies, $\{6, 10, 20, 50, 100, 200\}$ elements are used per dimension. We use Gauss rules in 1D (for the integration on the zero-level set Γ_0) and 2D (for the integration in Ω^-) with a rather high order of 11 which is kept constant independently of the order of background and interface elements. Results are studied in different error norms which are all related to integrating a function $f(\mathbf{x})$ either on Γ_0 or in Ω^-.

2.3.1 Integration on Zero-Isolines

Four different (relative) error norms are introduced for the integration on zero-level sets in two dimensions.

$$\varepsilon_1^\Gamma = \left|\left(\sum_i w_i\right) - I_1^\Gamma\right| / \left|I_1^\Gamma\right| \quad \text{with} \quad I_1^\Gamma = \int_{\Gamma_0} 1 \, \mathrm{d}\Gamma, \tag{6}$$

$$\varepsilon_f^\Gamma = \left|\left(\sum_i w_i \cdot f(\mathbf{x}_i)\right) - I_f^\Gamma\right| / \left|I_f^\Gamma\right| \quad \text{with} \quad I_f^\Gamma = \int_{\Gamma_0} f(\mathbf{x}) \, \mathrm{d}\Gamma, \tag{7}$$

$$\varepsilon_{f_{1D}^h}^\Gamma = \left|\left(\sum_i w_i \cdot f_{1D}^h(\mathbf{x}_i)\right) - I_f^\Gamma\right| / \left|I_f^\Gamma\right|, \tag{8}$$

$$\varepsilon_{f_{2D}^h}^\Gamma = \left|\left(\sum_i w_i \cdot f_{2D}^h(\mathbf{x}_i)\right) - I_f^\Gamma\right| / \left|I_f^\Gamma\right|, \tag{9}$$

where $\mathbf{x}_i$ are the integration points in the interface elements with integration weights w_i. The function $f(\mathbf{x})$ is defined as

$$f(\mathbf{x}) = {}^1\!/\!_2 \cdot x + {}^1\!/\!_4 \cdot y + x^2 + 2 \cdot y^3. \tag{10}$$

For $f^h_{1D}(\mathbf{x})$, $f(\mathbf{x})$ is prescribed at the nodes of the higher-order *interface* elements and interpolated inside the interface elements. Analogously, $f^h_{2D}(\mathbf{x})$ is the higher-order interpolation based on the *background* elements. For the error norm ε^Γ_f, the function (10) is evaluated *exactly* at the integration points $\mathbf{x}_i$. For the error norms $\varepsilon^\Gamma_{f^h_{1D}}$ and $\varepsilon^\Gamma_{f^h_{2D}}$, the function is interpolated at $\mathbf{x}_i$ by the higher-order shape functions in the interface or background elements, respectively. It is noted that, in contrast to the *nodes* of the interface elements, the integration points at $\mathbf{x}_i$ are only approximately on the zero-level set of $\phi^h(\mathbf{x})$.

Hence, $\phi^h(\mathbf{x})$ and $f^h(\mathbf{x})$ only approximate $\phi(\mathbf{x})$ and $f(\mathbf{x})$, respectively. Furthermore, the interface elements only approximate the zero-level set of $\phi^h(\mathbf{x})$. The numerical integration based on Gauss points only approximates the analytical integration. Nevertheless, although several approximations are involved, it is shown below that integration rules with optimal accuracy are recovered for integrands over $f(\mathbf{x})$ in domains defined by the level-set function $\phi(\mathbf{x})$.

Study 1: Variation of Search Paths

The deformed quadrilateral background mesh is used with Chen-Babuška nodes in the background and interface elements (case 1). The influence of the search paths as described in Sect. 2.1 is investigated. Results for ε^Γ_1 are shown in Fig. 11 for $\phi_2(\mathbf{x})$. Alternatives 1 and 2 were straight search paths based on the normal vector of the linear reconstruction or the gradient of the level-set function, respectively. Alternative 3 used the gradient at the intermediate positions throughout the iterative procedure to locate roots. It is seen that alternative 1 is slightly superior than 2 and 3. Because it is also particularly simple, it is used as a default in the remainder of this work.

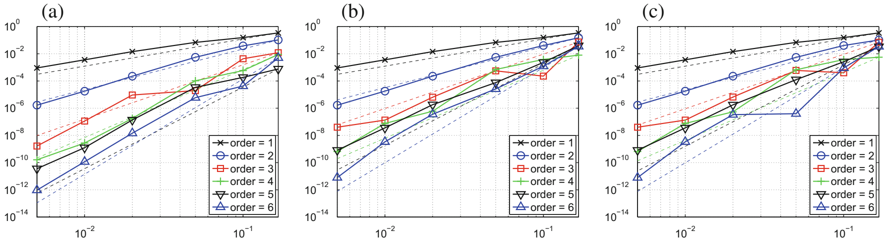


Fig. 11 Results for different search paths. (**a**) Alternative 1. (**b**) Alternative 2. (**c**) Alternative 3

Study 2: Variation of Error Norms

The different error norms are studied next, keeping the parameters of the previous study unchanged. The search directions are according to alternative 1. Results are seen in Fig. 12. It is noted that functions are integrated with optimal accuracy on the zero-level set, no matter whether exact function values are prescribed or interpolations in background or interface elements are used at the integration points.

Study 3: Variations of Elements

Above, results were obtained with deformed quadrilateral background meshes. The aim is now to show that equivalent results are obtained also for triangular elements and independently of deformed or undeformed background meshes, see Fig. 10. Results for ε_1^Γ are shown in Fig. 13 and confirm this assumption.

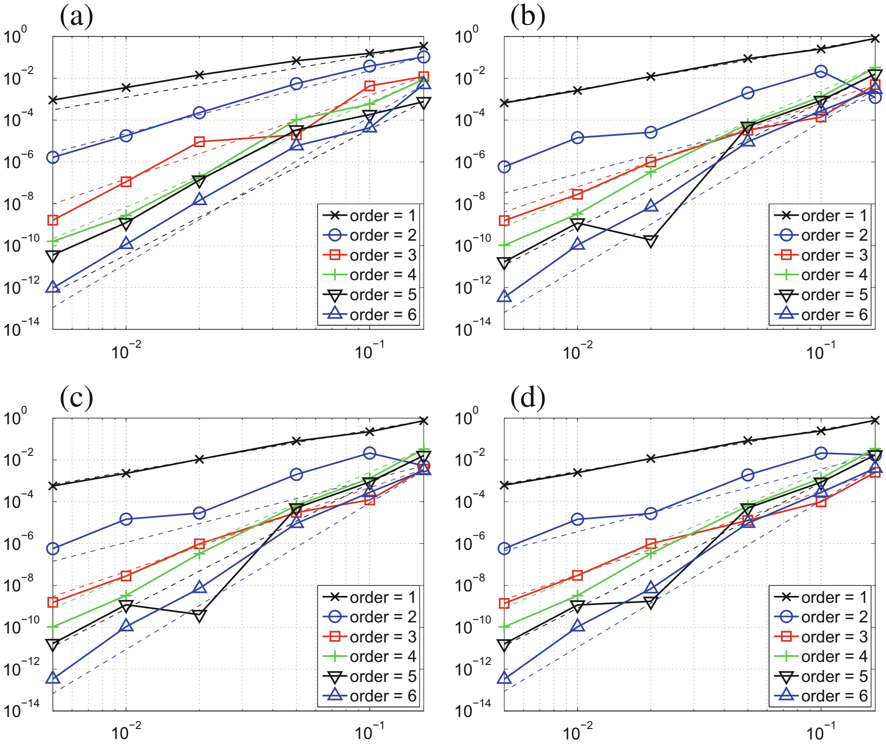


Fig. 12 Convergence in different error norms. (**a**) ε_1^Γ, (**b**) ε_f^Γ, (**c**) $\varepsilon_{f_{1D}^h}^\Gamma$, (**d**) $\varepsilon_{f_{2D}^h}^\Gamma$

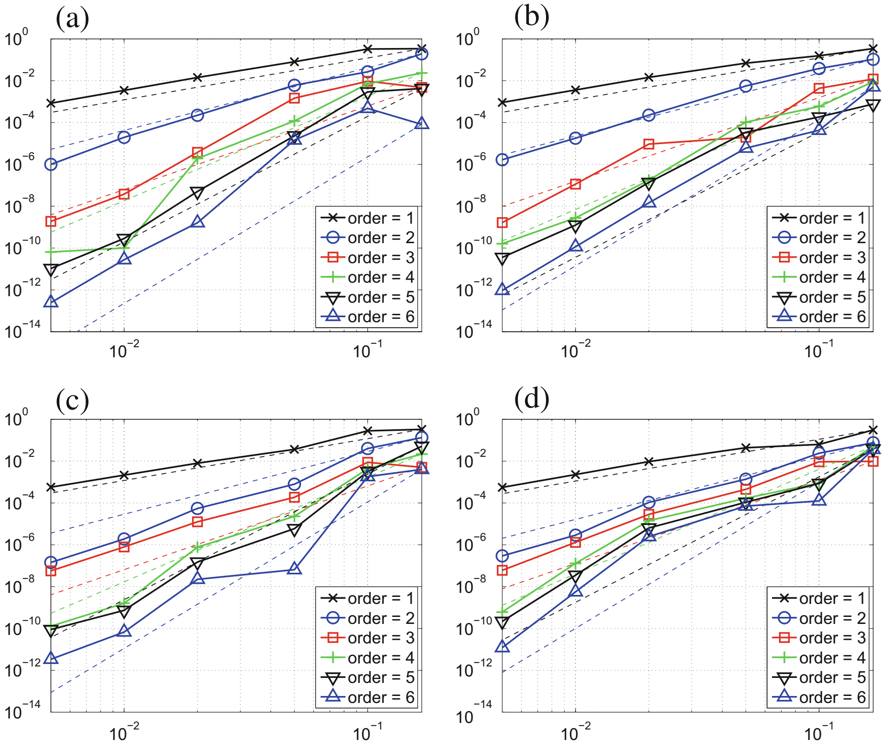


Fig. 13 Convergence for deformed/undeformed background meshes with triangular/quadrilateral elements. (**a**) Quad., undef., (**b**) quad., def., (**c**) tri., undef. (**d**) tri., def.

Study 4: Variation of Equally Spaced Nodes and Chen-Babuška Nodes

Again, we consider the deformed quadrilateral background mesh. All parameters are as in study 1 where the search paths are according to alternative 1. Equally spaced and Chen-Babuška nodes in the background and interface elements are systematically compared. The four possible combinations have already been described above. Results for ε_1^{Γ} are shown in Fig. 14. It is seen that the use of Chen-Babuška nodes yields significantly better results: the convergence rates are optimal in all cases yet the magnitude of the error varies with a factor of up to 10. For the best results, Chen-Babuška nodes are used in the background *and* interface elements.

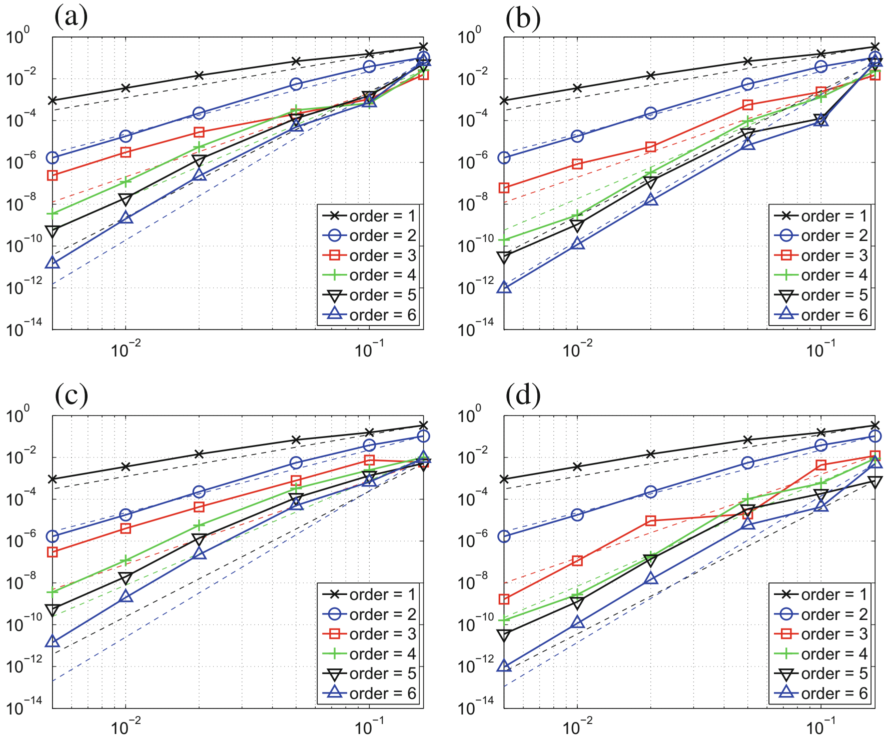


Fig. 14 Convergence in ε_1^Γ for equally spaced nodes or Chen-Babuška nodes in the background and interface elements, respectively. (**a**) Case 1: ES-ES. (**b**) Case 2: ES-CB. (**c**) Case 3: CB-ES. (**d**) Case 4: CB-CB

2.3.2 Integration in Cut 2D Elements

Three different (relative) error norms are introduced for the integration in Ω^-:

$$\varepsilon_1^\Omega = \left| \left(\sum_i w_i \right) - I_1^\Omega \right| / \left| I_1^\Omega \right| \quad \text{with} \quad I_1^\Omega = \int_{\Omega^-} 1 \, \mathrm{d}\Omega, \tag{11}$$

$$\varepsilon_f^\Omega = \left| \left(\sum_i w_i \cdot f(\mathbf{x}_i) \right) - I_f^\Omega \right| / \left| I_f^\Omega \right| \quad \text{with} \quad I_f^\Omega = \int_{\Omega^-} f(\mathbf{x}) \, \mathrm{d}\Omega, \tag{12}$$

$$\varepsilon_{fh}^\Omega = \left| \left(\sum_i w_i \cdot f^h(\mathbf{x}_i) \right) - I_f^\Omega \right| / \left| I_f^\Omega \right|, \tag{13}$$

where $\mathbf{x}_i$ are 2D integration points in the special sub-elements and w_i the corresponding weights. The standard setting from above is used: a deformed quadrilateral background mesh, search paths according to alternative 1, Chen-Babuška nodes in the background and interface elements, and the level-set function ϕ_2. Results in

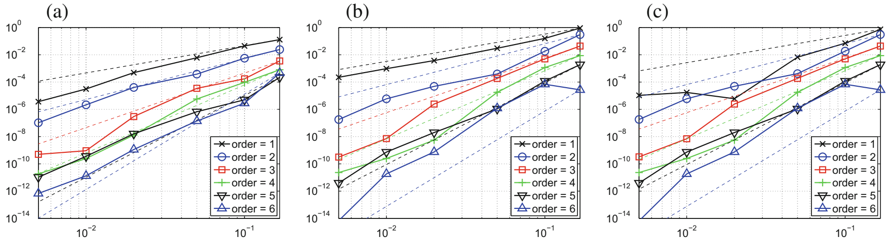


Fig. 15 Convergence in different error norms. (**a**) ε_1^{Ω}, (**b**) ε_f^{Ω}, (**c**) $\varepsilon_{fh}^{\Omega}$

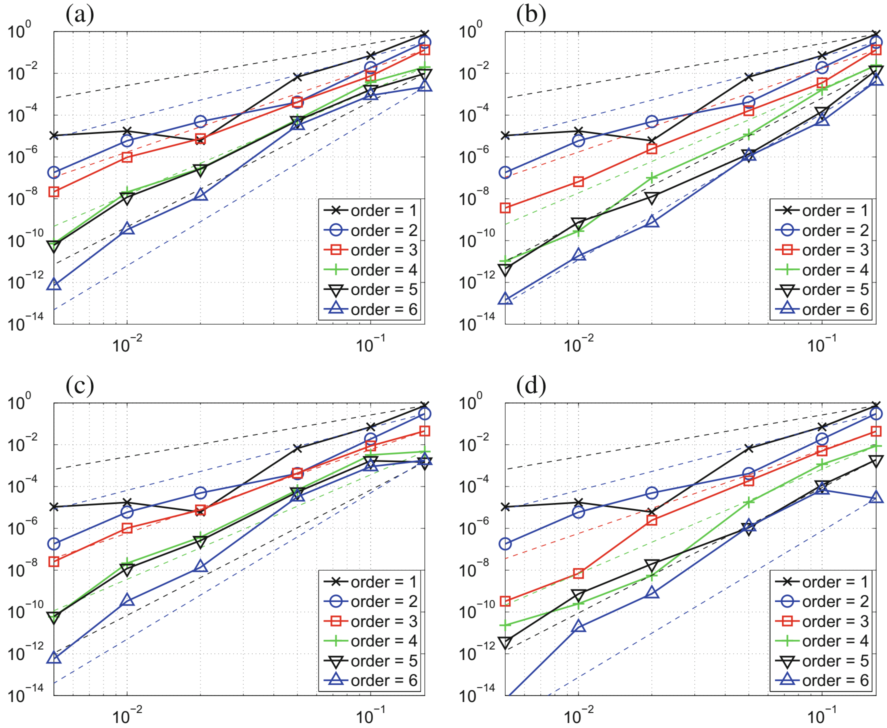


Fig. 16 Convergence in $\varepsilon_{fh}^{\Omega}$ for equally spaced nodes or Chen-Babuška nodes in the background and interface elements, respectively. (**a**) Case 1: ES-ES. (**b**) Case 2: ES-CB. (**c**) Case 3: CB-ES. (**d**) Case 4: CB-CB

the three different error norms are shown in Fig. 15. Optimal convergence rates are achieved in all cases. Next, equally spaced nodes and Chen-Babuška nodes are varied in the background and interface elements. Results are given in Fig. 16 and confirm the findings from above: Chen-Babuška nodes yield significantly better results than equally spaced nodes.

We obtained very similar results for different background elements in deformed and undeformed configurations. Also, ϕ_1 yields equivalent results so that convergence curves are not shown here for brevity. It is noted that for ϕ_1, I_1^{Ω} is the exact area of a circle and I_1^{Γ} the circumference.

3 Integration in Three Dimensions

The two-dimensional situation is now extended to the three-dimensional case. Hence, a three-dimensional higher-order background mesh built by Lagrange elements with equally-spaced nodes or Chen-Babuška nodes is considered. The element faces do not align with boundaries or interfaces, which are surfaces in 3D. In [5], Chen-Babuška nodes are given in tetrahedra, see Fig. 17a–c. For hexahedra as shown in Fig. 17d–f, the nodes are obtained by tensor-products from one-dimensional distributions as given in [4]. The shape functions associated to these elements feature the same properties as mentioned in Sect. 2.

The zero-level set Γ_0 of the level-set function $\phi(\mathbf{x})$, cf. Eq. (1), is a (curved) surface in three dimensions, typically defining boundaries or interfaces. As above,

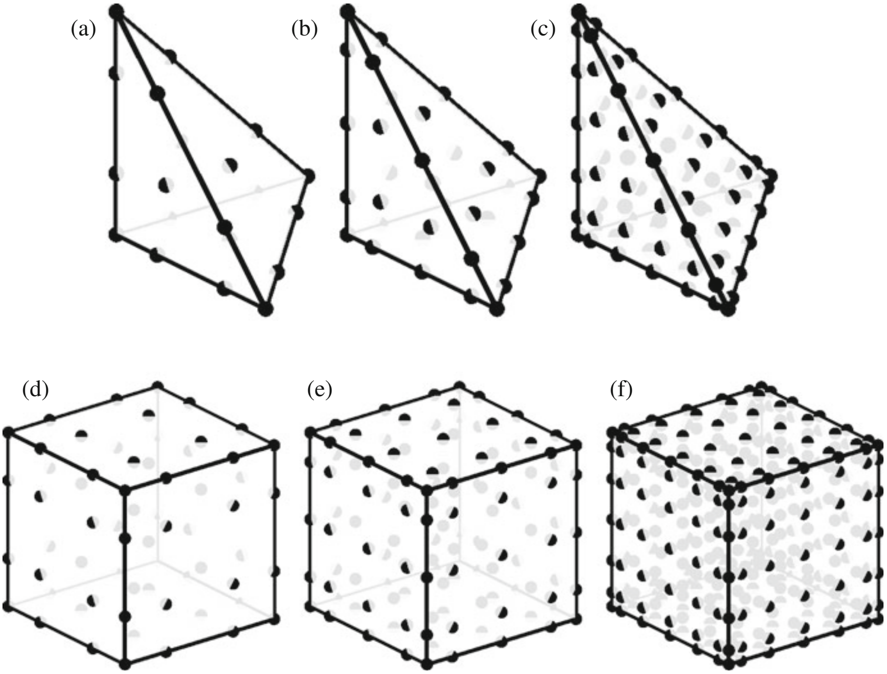


Fig. 17 Tetrahedral and hexahedral Chen-Babuška elements of different order. (**a**) Cubic, (**b**) 4th order, (**c**) 6th order, (**d**) cubic, (**e**) 4th order, (**f**) 6th order

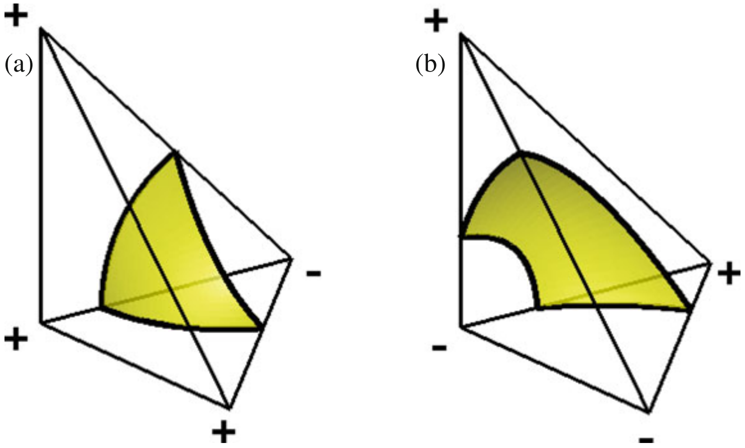


Fig. 18 Topologically different cut situations for valid level-set data in tetrahedra. (**a**) Topology 1, (**b**) topology 2

the level-set function is replaced by its interpolant $\phi^h(\mathbf{x})$, cf. Eq. (2). Again, the aim is to integrate on Γ_0 or inside the domains Ω^+ and/or Ω^-.

Starting point is the reconstruction of the zero-isosurface, i.e. the meshing by higher-order interface elements. This is first described for tetrahedra. We assume that valid level-set data is present, at least after some recursive refinements [10]. That is, the topological cut situation is defined based on the sign at the *corner* nodes of the tetrahedra. Then, only two different topologies are possible. Topology 1 is obtained when *three* edges are cut, the iso-surface is then a curved triangular element in the reference tetrahedron, and the two volumetric sub-elements are a sub-tetrahedron and a sub-prism. For topology 2, *four* elements edges are cut yielding a curved quadrilateral interface element and cutting the tetrahedron into two sub-prisms. The situation is shown in Fig. 18.

It is noted that hexahedral elements lead to a much larger number of topological cases even when only the signs at the corner nodes are considered. In a low-order context, these topologies are addressed by the marching cube algorithm [17], however, only for the detection and definition of the zero-level set. This algorithm does not address the decomposition into volumetric sub-elements for an integration on the two sides of the zero-level set. Therefore, it is advisable to first decompose a hexahedral element into tetrahedra and then proceed in each tetrahedron as mentioned below.

3.1 Reconstruction of Zero-Isosurfaces in 3D Background Elements

For the approximation of the zero-level set of $\phi^h(\mathbf{x})$ by higher-order interface elements, the corresponding element nodes have to be located on the zero-isosurface. The algorithm of the iterative procedure of Sect. 2.1, see Eq. 3, is straightforwardly extended to the three-dimensional case. The three alternatives for the search paths apply also in 3D.

An important difference, however, is the definition of starting points in the tetrahedra. It turns out that *linear* reconstructions, see Fig. 19a and d, do not provide good starting points for the iteration. The iteration often fails already for rather simple cases (modest curvature of the zero-level set in the tetrahedron) initiating (undesired) recursive refinements. It is much better to use a higher-order reconstruction based on *face* reconstructions of the tetrahedron. That is, for topology 1, a quadratic triangle element is generated by applying the 2D-reconstructions of Sect. 2.1 on the element faces, see Fig. 19b. For topology 3, a cubic Serendipity element is computed by face reconstructions, see Fig. 19e. Once these intermediate higher-order reconstructions are obtained, it is simple to map two-dimensional triangular or quadrilateral elements (with equally spaced nodes of Chen-Babuška nodes) onto them, see Fig. 19c and f. They provide excellent starting points for the root search. For the search paths, one may adapt the three alternatives for two-dimensional reconstructions, see Sect. 2.1. For alternative 1, it is noted that the normal vector at the starting points is no longer constant due to the curvature of the intermediate reconstruction. All alternatives converge quadratically towards a position on the zero-level set defining a node of the interface element.

For the integration on Γ_0, it is simple to map the obtained element nodes in the reference background element to the physical elements using the iso-parametric concept. Then, standard Gauss points in triangles and quadrilaterals are mapped to the physical interface elements. This follows the principle shown in two dimensions in Fig. 3. Gram's determinant is used to adjust the integration weights.

3.2 Integration in Cut 3D Background Elements

For the task to integrate in the three-dimensional domains Ω^+ and/or Ω^-, the cut elements have to be decomposed into volumetric sub-elements. As seen in Fig. 18, possible sub-elements are tetrahedra with one higher-order triangular side and prisms where the higher-order side is either a quadrilateral or a triangle. The higher-order side always matches with the reconstructed interface element. In [10], the mapping from the special (reference) sub-elements (where 3D Gauss points are used) to the cut reference background elements is defined based on the blending function mapping. These integration points are then mapped further to the physical background elements based on the isoparametric concept. The issue of "consistent

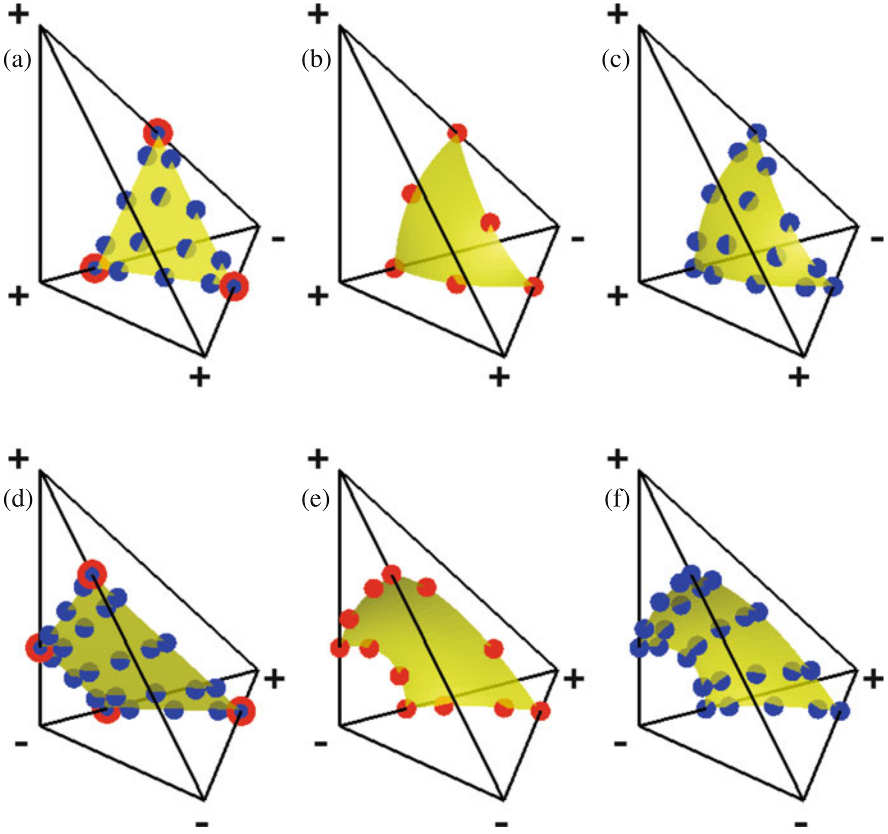


Fig. 19 Reconstruction in tetrahedra. (**a**) and (**d**) show starting points based on a *linear* reconstruction which turned out to be less robust. Therefore, for topology 1, a *quadratic* triangle and for topology 2, a *cubic* Serendipity element are defined by 2D-reconstructions on the faces, see (**b**) and (**e**). The desired higher-order element is mapped onto these interface elements in order to provide (highly accurate) starting points for the root finding, see (**c**) and (**f**). (**a**) Top. 1, linear. (**b**) Top. 1, face recon. (**c**) Top. 1, starting points. (**d**) Top. 2, linear. (**e**) Top. 2, face recon. (**f**) Top. 2, starting points

integration rules" which consider the special properties of the mapping is pointed out in [10]; it restricts the selection of possible Gauss rules in the (reference) sub-elements.

Note that the blending function mapping may lead to negative Jacobians even if the reconstruction of the interface element, defining the higher-order side of the sub-elements, was successful. The Jacobians must therefore be checked and a recursive refinement may be necessary to obtain valid mappings. We find the following reasons for recursive refinements: (1) to obtain "valid" level-set data, (2) to provide suitable roots inside the reference background element, and (3) to obtain non-negative Jacobians for the blending function mapping.

3.3 Numerical Results in Three Dimensions

Numerical results are now presented for the integration of three-dimensional geometries. First for integrations on $\varGamma_0$, then for integrations in $\varOmega^-$. The procedure is similar to the studies in two dimensions, see Sect. 2.3. For the search paths of the nodes of the interface elements, we restrict ourselves to the one which is based on the normal vector of the intermediate reconstruction. Note that in 3D, the intermediate reconstruction is already higher-order (quadratic triangle elements or cubic Serendipity elements) so that the normal vectors are not constant for all starting points as in 2D. The gradient-based alternatives were leading to worse results and are, therefore, omitted. We consider the following two level-set functions in our studies:

$$\phi_1(\mathbf{x}) = \sqrt{x^2 + y^2 + z^2} - r \tag{14}$$

$$\phi_2(\mathbf{x}) = \phi_1(\mathbf{x}) + 0.1 \cdot [\cos(2\pi \cdot x) + \cos(2\pi \cdot y) + \cos(2\pi \cdot z)] \tag{15}$$

with $r = 0.7123$. The zero-isosurfaces of these two level-set functions are visualized in Fig. 20. There, also the undeformed Cartesian background mesh is shown, deformed configurations in three dimensions are neglected here. For the integrand, the function

$$f(\mathbf{x}) = x^2 + y^2 + {}^1\!/_2 \cdot \cos(z) \tag{16}$$

is chosen.

Hexahedral and tetrahedral background meshes are studied. In the convergence studies, $\{2, 4, 6, 10, 14, 20, 30, 50, 100\}$ elements are used per dimension. Of course, especially for the coarse meshes recursive refinements are necessary in a larger percentage of the elements than in the fine meshes. These refinements are not considered for the element spacing h in the convergence plots shown herein, leading to some wiggles in the convergence behavior.

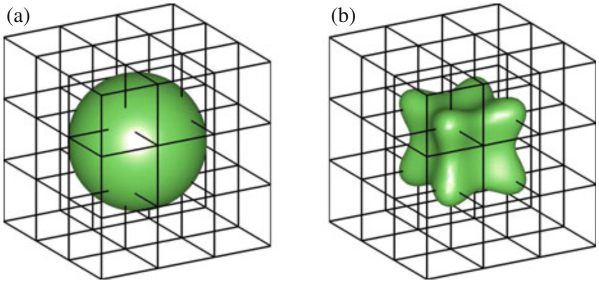


Fig. 20 Zero-isosurfaces of ϕ_1 and ϕ_2 in $[-1, 1]^3$ according to Eqs. (14) and (15), respectively. (**a**) Zero-isosurface of ϕ_1, (**b**) zero-isosurface of ϕ_2

3.3.1 Integration on Zero-Isosurfaces

The error norms of Eqs. (6)–(9) are straightforwardly extended to the three-dimensional case, leading to ε_1^Γ, ε_f^Γ, $\varepsilon_{f_{2D}^h}^\Gamma$, $\varepsilon_{f_{3D}^h}^\Gamma$ which are integrated on the zero-isosurface Γ_0. Results are shown in Fig. 21 for hexahedral background meshes and ϕ_2. Comparing the upper and the lower figures of Fig. 21, respectively, shows that Chen-Babuška points in the elements again yield to better results than equally spaced nodes. Furthermore, comparing the left and right figures in Fig. 21, respectively, shows that the error converges similar in all error norms. We only show a representative subset of the results for brevity. Note that, most importantly, the convergence rates are higher-order yet not optimal in three dimensions. This is in agreement with [10]. There, this issue was further traced back to the fact that the boundaries of the interface elements are strictly enforced to be on the element faces of the background elements.

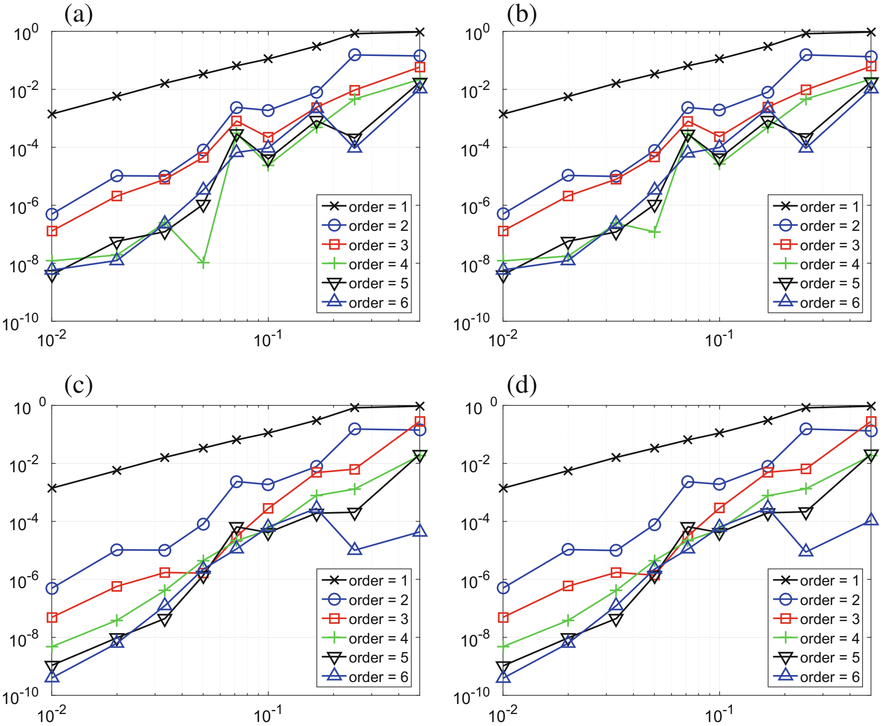


Fig. 21 Convergence for the integration on Γ_0 in three dimensions. (**a**) Case 1: ES-ES, ε_1^Γ. (**b**) Case 1: ES-ES, $\varepsilon_{f_{3D}^h}^\Gamma$. (**c**) Case 4: CB-CB, ε_1^Γ. (**d**) Case 4: CB-CB, $\varepsilon_{f_{3D}^h}^\Gamma$

3.3.2 Integration in Cut 3D Elements

The error norms of Eqs. (11)–(13) are straightforwardly extended to the three-dimensional case. However, as seen for the integration on the zero-level sets, the convergence behavior is independent of these norms so that results are only shown for $\varepsilon_{fh}^{\Omega}$. Furthermore, we restrict ourselves to the case of ϕ_1, where the integration is performed in a ball with radius r. The advantage is that there are analytical results available for the integration so that numerical reference solutions are not needed. For ε_1^{Ω} follows the volume of a ball, $I_1^{\Omega} = 4/3 \cdot \pi \cdot r^3$, and for ε_f^{Ω} and $\varepsilon_{fh}^{\Omega}$ follows $I_f^{\Omega} = 2 \cdot \pi \cdot (\sin(r) - r \cdot \cos(r)) + 4/105 \cdot \pi \cdot r^7$ when f is defined by Eq. (16). Results are seen in Fig. 22 and confirm the findings from above.

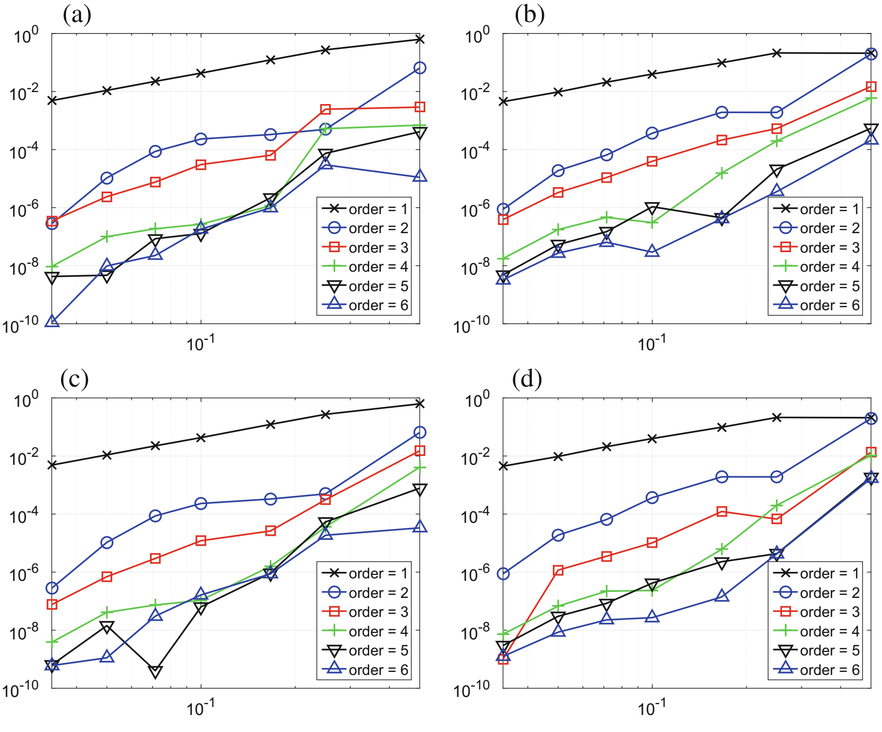


Fig. 22 Convergence in $\varepsilon_{fh}^{\Omega}$ for the integration in Ω^{-} in three dimensions. (**a**) Case 1: ES-ES, hexa. (**b**) Case 1: ES-ES, tetra. (**c**) Case 4: CB-CB, hexa. (**d**) Case 4: CB-CB, tetra

4 Conclusions

The use of special element nodes as proposed by Chen and Babuška in [4, 5] in the context of higher-order accurate integrations of implicit geometries in two and three spatial dimensions is investigated. Two standard integration tasks are studied, the integration on zero-level sets and in the sub-domains implied by the sign of the level-set function. In a first step, the algorithm relies on the reconstruction of the zero-level set of the interpolated level-set function in terms of higher-order interface elements. In the second step, sub-elements are defined which feature a curved side which coincides with the higher-order interface element. The blending function mapping is used to map standard Gauss points into these special sub-elements.

The underlying algorithm was introduced and described in detail in [10]. Herein, it is extended to the use of Chen-Babuška elements. It is found that this improves the error by a factor of typically 5 to 10 for the same number of elements, i.e. the same computational effort. Curved geometries are studied in two and three dimensions and convergence studies are made for (1) different element types, (2) Cartesian or perturbed background meshes, and (3) different search directions. Optimal results have been obtained in two dimensions independently of the element type and state of the background meshes. The search directions based on the normal vector of the linear reconstructions of the zero-level set performs better than those based on the gradient of the level-set function. In three dimensions, the results are higher-order accurate yet sub-optimal. In [10], this was at least partly traced back to the fact that the sub-elements are strictly enforced to be inside the background reference elements. The errors are checked in different norms. Some error norms also involve interpolated integrands instead of exact values at the integration points and it is found that this does not reduce the accuracy.

The proposed integration scheme allows for a consistent integration in higher-order accurate numerical methods, e.g. in the frame of fictitious domain methods and extended finite element methods.

References

1. Abedian, A., Parvizian, J., Düster, A., Khademyzadeh, H., Rank, E.: Performance of different integration schemes in facing discontinuities in the finite cell method. Int. J. Comput. Methods **10**, 1–24 (2013)
2. Belytschko, T., Black, T.: Elastic crack growth in finite elements with minimal remeshing. Int. J. Numer. Methods Eng. **45**, 601–620 (1999)
3. Burman, E., Hansbo, P.: Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method. Comput. Methods Appl. Mech. Eng. **199**, 2680–2686 (2010)
4. Chen, Q., Babuška, I.: Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. Comput. Methods Appl. Mech. Eng. **128**, 405–417 (1995)

5. Chen, Q., Babuška, I.: The optimal symmetrical points for polynomial interpolation of real functions in the tetrahedron. Comput. Methods Appl. Mech. Eng. **137**, 89–94 (1996)
6. Cheng, K., Fries, T.: Higher-order XFEM for curved strong and weak discontinuities. Int. J. Numer. Methods Eng. **82**, 564–590 (2010)
7. Dréau, K., Chevaugeon, N., Moës, N.: Studied X-FEM enrichment to handle material interfaces with higher order finite element. Comput. Methods Appl. Mech. Eng. **199**, 1922–1936 (2010)
8. Düster, A., Parvizian, J., Yang, Z., Rank, E.: The finite cell method for three-dimensional problems of solid mechanics. Comput. Methods Appl. Mech. Eng. **197**, 3768–3782 (2008)
9. Fries, T., Belytschko, T.: The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Methods Eng. **84**, 253–304 (2010)
10. Fries, T., Omerović, S.: Higher-order accurate integration of implicit geometries. Int. J. Numer. Methods Eng. doi:10.1002/nme.5121 (2015, accepted)
11. Gordon, W., Hall, C.: Construction of curvi-linear co-ordinate systems and applications to mesh generation. Int. J. Numer. Methods Eng. **7**, 461–477 (1973)
12. Gordon, W., Hall, C.: Transfinite element methods: blending function interpolation over arbitrary curved element domains. Numer. Math. **21**, 109–129 (1973)
13. Laborde, P., Pommier, J., Renard, Y., Salaün, M.: High-order extended finite element method for cracked domains. Int. J. Numer. Methods Eng. **64**, 354–381 (2005)
14. Legay, A., Wang, H., Belytschko, T.: Strong and weak arbitrary discontinuities in spectral finite elements. Int. J. Numer. Methods Eng. **64**, 991–1008 (2005)
15. Legrain, G., Chevaugeon, N., Dréau, K.: High order X-FEM and levelsets for complex microstructures: uncoupling geometry and approximation. Comput. Methods Appl. Mech. Eng. **241–244**, 172–189 (2012)
16. Li, Z.: A fast iterative algorithm for elliptic interface problems. SIAM J. Numer. Anal. **35**, 230–254 (1998)
17. Lorensen, W., Cline, H.E.: Marching cubes: a high resolution 3d surface construction algorithm. SIGGRAPH Comput. Graph. **21**, 163–169 (1987)
18. Mittal, R., Iaccarino, G.: Immersed boundary methods. Annu. Rev. Fluid Mech. **37**, 239–261 (2005)
19. Moës, N., Dolbow, J., Belytschko, T.: A finite element method for crack growth without remeshing. Internat. J. Numer. Methods Eng. **46**, 131–150 (1999)
20. Moumnassi, M., Belouettar, S., Béchet, É., Bordas, S., Quoirin, D., Potier-Ferry, M.: Finite element analysis on implicitly defined domains: an accurate representation based on arbitrary parametric surfaces. Comput. Methods Appl. Mech. Eng. **200**, 774–796 (2011)
21. Mousavi, S., Sukumar, N.: Numerical integration of polynomials and discontinuous functions on irregular convex polygons and polyhedrons. Comput. Mech. **47**, 535–554 (2011)
22. Müller, B., Kummer, F., Oberlack, M.: Highly accurate surface and volume integration on implicit domains by means of moment-fitting. Internat. J. Numer. Methods Eng. **96**, 512–528 (2013)
23. Osher, S., Fedkiw, R.: Level Set Methods and Dynamic Implicit Surfaces. Springer, Berlin (2003)
24. Phillips, G.: Interpolation of Approximation by Polynomials. Springer, Berlin (2013)
25. Rivlin, T.: An Introduction to the Approximation of Functions. Dover, New York, NY (1981)
26. Sala-Lardies, E., Fernández-Méndez, S., Huerta, A.: Optimally convergent high-order X-FEM for problems with voids and inclusion. In: Proceedings of the ECCOMAS 2012, Vienna, Austria (2012)
27. Sethian, J.: Level Set Methods and Fast Marching Methods, 2nd edn. Cambridge University Press, Cambridge (1999)
28. Stazi, F., Budyn, E., Chessa, J., Belytschko, T.: An extended finite element method with higher-order elements for curved cracks. Comput. Mech. **31**, 38–48 (2003)
29. Szabados, J., Vertesi, P.: Interpolation of Functions. World Scientific, Singapore (1990)
30. Szabó, B., Düster, A., Rank, E.: The p-Version of the Finite Element Method, chap. 5, pp. 119–139. Wiley, Chichester (2004)

31. Ventura, G.: On the elimination of quadrature subcells for discontinuous functions in the eXtended finite element method. Int. J. Numer. Methods Eng. **66**, 761–795 (2006)
32. Ventura, G., Gracie, R., Belytschko, T.: Fast integration and weight function blending in the extended finite element method. Int. J. Numer. Methods Eng. **77**, 1–29 (2009)
33. Zi, G., Belytschko, T.: New crack-tip elements for XFEM and applications to cohesive cracks. Int. J. Numer. Methods Eng. **57**, 2221–2240 (2003)