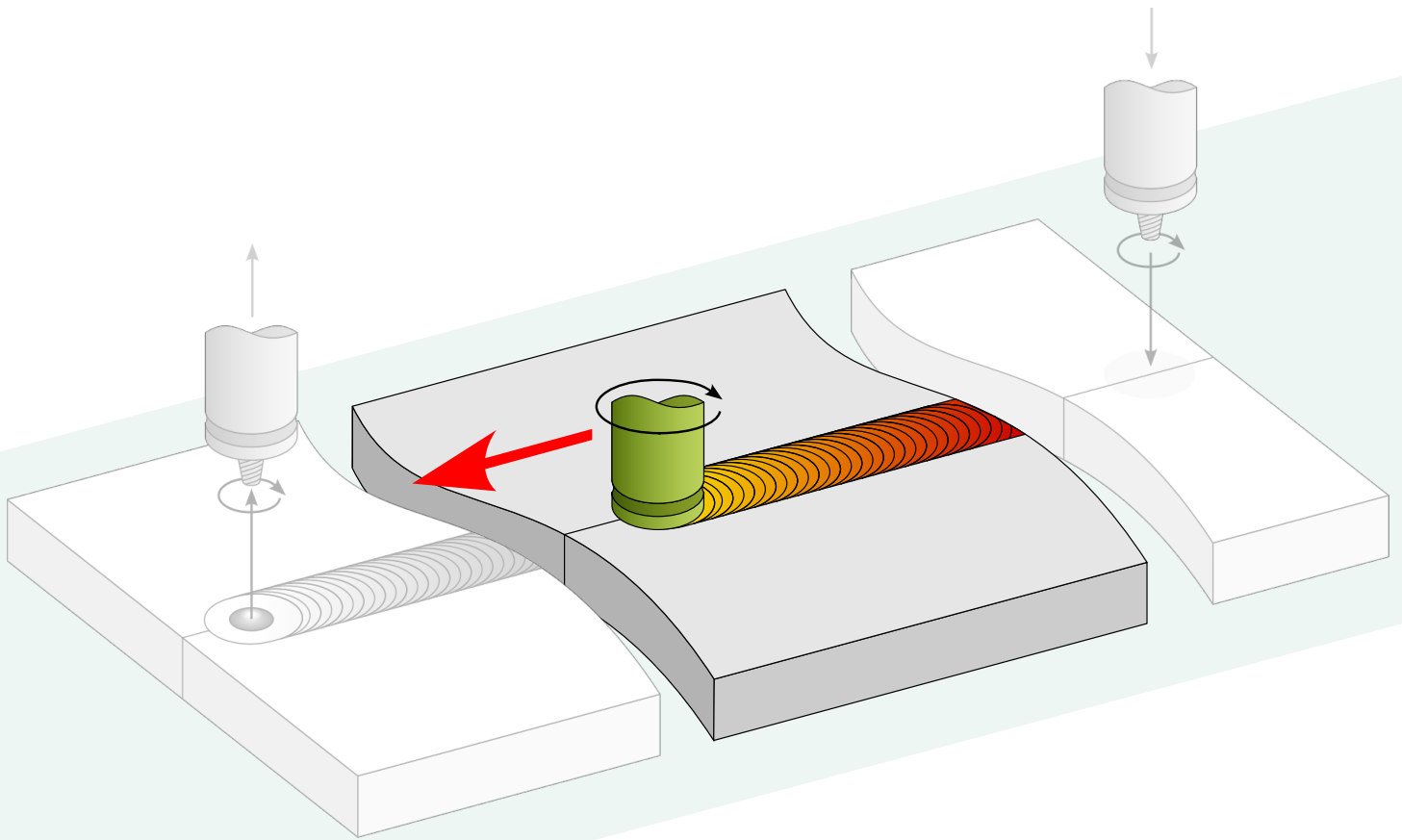


Advanced Finite Element Methods for Metal Forming and Manufacturing Process Simulation: An Application to Friction Stir Welding Analysis.

Month 2025



PhD. Thesis by **Henning Venghaus**

Supervisors	Prof. Michele Chiumenti
	Prof. Joan Baiges Aznar
Tutor	Prof. Daniel Juhre



Contents

Contents	ii
1 Introduction	2
1.1 Motivation	2
1.2 Objectives	5
1.3 Outline	6
2 Element Technology	8
2.1 Objectives	8
2.1.1 Isochoric Material Behavior	8
2.1.2 Strain & Stress Accuracy	11
2.2 State Of The Art	11
2.2.1 h , p and hp -refinement	12
2.2.2 Reduced and Selective Reduced Integration	12
2.2.3 $\bar{\mathbf{B}}$ -Methods	13
2.2.4 Enhanced Assumed Strain (EAS)	15
2.2.5 $\mathbf{Q}_1\mathbf{P}_0$	16
2.3 Mixed Formulations	16
2.3.1 Two-field u/p Formulation	18
2.3.2 Three-field $u/p/e$ Formulation	20
2.4 Numerical Stability and Stabilization	24
2.5 Adaptive Formulation Refinement	29
2.5.1 Motivation	29
2.5.2 Methodology	31
2.5.3 Combination with h -refinement	32
2.5.4 Refinement Criteria	32
2.5.5 Solver Strategies	34
2.5.6 Results	36
2.6 Concluding Remarks	39
3 Unfitted FEM / Embedded Domain Methods	40
3.1 Objectives	40
3.2 State Of The Art	42
3.2.1 Geometry Description and Element Classification	42
3.2.2 Integration Method	45
3.2.3 Boundary Condition Application	49
3.2.4 Small Cut Stabilization	50
3.2.5 Enrichment	52
3.2.6 Established Methods	54
3.3 Mesh Refinement and Multi-Refinement Criteria	55
3.4 Concluding Remarks	57
4 Friction Stir Welding	59
4.1 State of the Art FSW Technology	59
4.2 State of the Art FSW Simulations	71
4.3 Finite Element Framework	78
4.3.1 Frame of Reference	79
4.3.2 Thermo-Mechanical Coupling	79
4.3.3 Boundary Conditions	81
4.3.4 Finite Element Approximation	83

4.3.5	Concave Tools	86
4.3.6	Surface Normals	87
4.4	Numerical Results	88
4.5	Bridging Research and Industrial Application	90
4.6	Concluding Remarks	96
5	Conclusions	98
5.1	Summary	98
5.2	Contribution	99
5.3	Future Work	101
	Bibliography	103
A	Publication 1	116
B	Publication 2	139
C	Publication 3	163

List of Figures

1.1	Metal forming processes	2
1.2	Emerging metal manufacturing techniques	4
2.1	Examples of the occurrence of isochoric flow and incompressible materials	8
2.2	Mesh dependent crack evolution	11
2.3	Refinement approaches	12
2.4	Hourglassing in a cylinder under internal pressure due to reduced integration.	13
2.5	Element types and native <i>inf-sup</i> satisfaction	24
2.6	Global error of displacements, pressure and deviatoric strain over element length h and degrees of freedom.	30
2.7	Displacement and strain interpolation at AFR interface.	31
2.8	Four elements with potentially 7 DOF per node	35
2.9	Four elements with potentially 7 DOF per node	35
2.10	Global error of displacements (left) and deviatoric strain (right) over element length h (top) and degrees of freedom (bottom)	37
2.12	Crack evolution in branching test at different stages	38
3.2	Mesh generation from geometry: body fitted and non-body fitted	40
3.3	Element Classification	43
3.4	Selection of domain description methods	43
3.5	Use of Level-set Functions for the Description of an Inclusion (a) and a Crack (c-d)	44
3.6	Integration Rule for boundary-conform integration of two triangle elements.	45
3.8	Octree refinement of sub-integration cells.	46
3.7	Decomposition of a quadrilateral into two triangles and a hexahedron into five tetrahedra.	46
3.9	Volume and Surface Integration Rules obtained with Moment Fitting.	48
3.10	Cut elements and two layers of physical elements, stabilized using the Ghost Penalty method.	52
3.11	Local enrichment of approximation space	53
3.12	Refinement levels and layers	55
3.13	Balancing of refinement levels	56
3.14	Workflow of Embedding CAD Geometry into Computational Domain, Computing Level-Set Function and Classifying Elements	56
4.1	Friction Stir Welding Process	59
4.2	Completed, sound weld.	61
4.3	Different zones obtained after FSW	61
4.4	FSW machines	62
4.5	Pin-tool	63
4.6	Shoulder Types	64
4.7	Pin Types	64
4.8	Possible weld defects due to displacement controlled setup (right), compared to force controlled process (left).	65
4.9	Application of FSW transportation.	69
4.10	FSW mesh detail	71
4.11	Lagrangian framework	73
4.12	Eulerian framework	74
4.13	ALE framework	75

4.14 AKF framework	76
4.15 AKF framework	76
4.16 EDM framework	77
4.17 Evolution of analytical and numerical methods to analyze FSW processes	78
4.18 Overview FSW Frameworks	78
4.19 Tool advancement (left) vs. material inflow modeling (right)	79
4.20 Thermo-mechanical coupling	80
4.21 Temperature and strain rate dependent effective viscosity and equivalent stress	81
4.22 Linearized material law	81
4.23 Thermal and mechanical boundary conditions on the exterior boundaries (left) and the tool-workpiece interface Γ_I , visualized as a cross section of the domain (right)	82
4.24 Kinematics and exterior normals on the tool-workpiece interface Γ_I : cross section (left), top view for axisymmetric pin (middle), and top view for non-axisymmetric pin (right)	82
4.25 2D slice of computational domain. Level set function highlighted in red.	85
4.26 Harmonic extension and one-sided domain decoupling	85
4.28 Different approaches to handling concave tools	86
4.29 Excluded elements	87
4.31 Recovery of exact normals	88
4.32 Definition of torque and advancing, lateral and plunging reaction forces	88
4.33 A set of different tools: cylindrical (left), threaded (middle) and trivex (right)	89
4.34 Forces, torque and maximum Temperature evolution over time	89
4.35 Impact of tilting angle ($0^\circ \rightarrow 2^\circ$) on reaction forces, torque and maximum temperature	90
4.36 Graphical User Interface	92
4.37 Temperature measurements in Lagrangian framework (material temperature, left) and Eulerian framework (spacial temperature, right)	93
4.38 Particle Tracing	95

List of Tables

2.1	Poisson's Ratios of common Metals	8
2.2	Local and global convergence rates for the primary unknown and a 1 st order derivative . . .	11
2.3	Different stabilization approaches	26
2.4	Global and local convergence rates for displacement, pressure and strain for different stabilization approaches (linear/linear/linear elements)	26
2.5	Number of degrees of freedom and size of elemental stiffness matrices for tetrahedral and hexahedral elements.	29
2.6	Refinement and de-refinement thresholds	38
4.1	List of most important FSW simulation parameters	91

Acronyms

In this section a list of several acronyms which are used along the thesis is stated.

ALE	Arbitrary Lagrangian-Eulerian
AKF	Apropos Kinematic Framework
ASGS	Algebraic Sub-Grid Stabilization
BREP	Boundary Representation
CAD	Computer-Aided Design
CutFEM	Cut Finite Element Method
DG	Discontinuous Galerkin
EDM	Embedded Domain Method
FE	Finite Element
FEM	Finite Element Method
FSP	Friction Stir Processing
FSW	Friction Stir Welding
GFEM	Generalized Finite Element Method
GS	Ghost Stabilization
IBM	Immersed Boundary Method
IGES	Initial Graphics Exchange Specification
OSS	Orthogonal Subscale Stabilization
PUM	Partition of Unity Method
SBM	Shifted Boundary Method
STL	STereoLithography
STEP	STandard for the Exchange of Product model data
VMS	Variational Multiscale Stabilization
XFEM	eXtended Finite Element Method

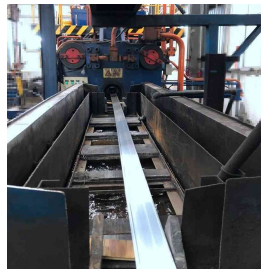
1.1 Motivation

Metal forming processes A cornerstone of today's engineering achievements are, without question, metal forming processes, spanning all application areas and offering a multitude of possibilities. In civil engineering, metal forming processes such as rolling and forging play a pivotal role in producing structural beams and columns. These components serve as the structural foundation of modern infrastructure, supporting the development of urban landscapes through essential elements like buildings, bridges, and tunnels. Extrusion and stretching processes are used to manufacture reinforcement bars (rebar), which are essential in virtually any reinforced concrete structure. Extrusion is used to manufacture freshwater and sewage pipes and oil and gas pipelines.

In mechanical engineering, forging, and extrusion techniques are used to produce blanks of machine components such as gears, shafts, and bearings before machining is applied to obtain the final dimensions and desired surface finishes. Roll-forming is used to manufacture a variety of cross-section profiles with excellent attributes, such as low weights due to thin walls and high strength due to being work-hardened. Die forming is used to manufacture various components, including consumer goods such as beverage cans, casings for electronic devices or household items such as washing machines, and kitchenware like pots and pans.



(a) Rail wheel forging ^[a]



(b) Extrusion ^[b]



(c) Hot-rolled rails ^[c]

In transportation engineering, rolling and forging are used to manufacture train wheels (Fig. 1.1a), rails (Fig. 1.1c), and fasteners. Modern train carriages are often manufactured from extruded aluminum profiles (Fig. 1.1b), and car bodies are made from deep-drawn metal sheets. Global commerce as it exists today relies heavily on cargo ships and shipping containers, both of which are constructed from rolled steel plates and sheets, respectively.

Metal forming processes are typically characterized by plastic and rate-dependent deformations, often involving large strains and, depending on the specific process, also high strain rates. The majority of metals

Figure 1.1: Metal forming processes

^a <https://rail-wheels.com/...>
(cropped, accessed: 25.10.2024)

^b <https://sourceint.com/...>
(cropped, accessed: 25.10.2024)

^c <https://metallonline.ch/...>
(accessed: 25.10.2024)

and alloys used in modern engineering applications (such as Steel, Aluminum, or Titanium) exhibit J_2 plasticity, meaning that the plastic behavior is independent of the hydrostatic pressure and is solely driven by deviatoric deformations. It is well known that materials that experience softening exhibit strain localizations when subjected to monotonic straining. In particular, the aforementioned J_2 materials form shear bands under the right conditions [1].

A considerable amount of metal processing is performed at elevated temperatures. Either because the workpieces are intentionally preheated above specific material-dependent temperatures, which facilitates recrystallization, for example, in hot-rolling, forging, and hot extrusion, or because the plastic deformations dissipate enough energy, as in cold-rolling. As a result, the workpiece heats up, and thermal strains (although smaller than the mechanical strains) might not always be negligible [2]. Either way, thermal expansion and temperature-dependent material properties such as Young's Modulus, yield strength, and plastic flow rates all influence the forming processes, and the obtained grain structures have an enormous effect on the final components' mechanical properties.

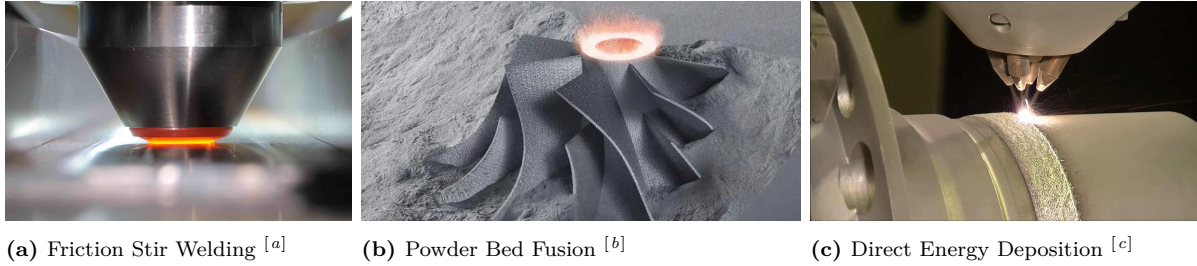
Manufacturing Processes Metal forming processes only represent half of the equation. Many of the machines, devices, and structures we rely on daily – from consumer goods such as kitchen appliances to train carriages and spacecraft – are built using stock materials processed through the above-mentioned metal forming processes. However, creating fully assembled products from these components often requires joining techniques to bring together various parts and materials into a cohesive whole. Next to riveting, which is somewhat outdated (except for aerospace applications), and soldering, brazing, and gluing, which are typically not used for load-bearing structures, welding stands out as the most common and effective joining technique for metals.

Welding is a thermo-mechanical joining process that melts parts together, causing fusion. Typically a filler material is also added. The process induces phase transformations in the material and substantial thermal strains. During welding, the melt pool exhibits isochoric flow, and residual stresses can appear that might impede the use of the manufactured component. Ultimately, it shows very similar characteristics to the above-mentioned metal forming processes.

Even comparatively novel manufacturing techniques share most of the same characteristics. Friction Stir Welding (FSW, Fig. 1.2a) is a solid-state joining process where two workpieces are welded together using a rotating pin tool. FSW can have significant advantages over conventional welding techniques but is ultimately governed by the same aspects: strong thermo-mechanical coupling, plastic deformations, isochoric material flow, and large strains and strain rates. Even (metal) additive manufacturing techniques like power bed fusion (PBF, Fig. 1.2b) or direct energy deposition (DED, Fig. 1.2c) share the same profile, although the occurring strains are typically much lower than in FSW.

[1]: Cervera et al. (2004), “Shear band localization via local J_2 continuum damage mechanics”

[2]: Koohbor et al. (2011), “Thermo-mechanical behaviours of strip and work-rolls in cold rolling process”



Numerical Methods For any manufacturing technique, researchers and engineers have conducted studies and experiments to determine the best methods and processing parameters to obtain the most satisfying results. While a lot of knowledge about a particular method is gained through experiments, numerical simulations (and the Finite Element Method in particular) are vital tools to accelerate this process and deepen our understanding of it.

On the one hand, numerical simulations can visualize material flow patterns, temperature distributions, strains, and stresses in ways that are impossible to visualize by examining physical specimens. On the other hand, simulations are typically cheaper, less time-consuming and can be performed in parallel, testing hundreds of parameter combinations to obtain an ideal extrusion ratio, the fastest possible welding speeds or optimize grain sizes. In essence, simulations allow for the digital design and optimization of products and manufacturing methods, therefore reducing development costs and time and helping to use available resources more economically.

The critical aspects of any numerical simulations are *accuracy*, *efficiency*, and *robustness*. Accuracy must always be considered in the context of the complexity of the physical problem at hand. Metal forming and manufacturing applications pose some unique challenges. The large strains and strain rates, respectively, need to be addressed, as well as the isochoric nature of the plastic deformations. The thermo-mechanical coupling poses a further challenge, and novel manufacturing techniques such as FSW and AM introduce complicated geometries and boundary conditions. The former is characterized by enormous strain rates close to the complex tool geometry and the latter allows to precisely manufacture the most complex geometries.

Computational efficiency is essential to achieve accurate simulation results quickly and with manageable resource use. This can involve developing adaptive refinement algorithms to concentrate computational efforts on key areas while reducing mesh density, and thus cost, in less critical regions. Alternatively, the numerical model can be tailored to address specific questions rather than simulating an entire manufacturing process from start to finish.

Finally, robustness maximizes the chance of a numerical method converging to a solution and ensures that the numerical method itself depends as little as possible on problem-specific parameters.

Technology Transfer The significance of numerical simulations is not only known to researchers and manufacturers but also law makers.

Figure 1.2: Emerging metal manufacturing techniques

^a [https://twi-global.com/...](https://twi-global.com/)
(accessed: 25.10.2024)

^b [https://us.ark.com/...](https://us.ark.com/)
(accessed: 25.10.2024)

^c [https://pick3dprinter.com/...](https://pick3dprinter.com/)
(accessed: 25.10.2024)

Public-private partnerships such as the European High-Performance Computing Joint Undertaking (EuroHPC JU) have gained traction and enabled the pooling of resources across all of Europe to ...

“[...] support the development of a European supercomputing ecosystem, stimulating a technology supply industry, and making supercomputing resources in many application areas available to a large number of public and private users, including small and medium-sized enterprises.” [3]

The undertaking is jointly funded by its members and has a budget of around € 7 billion from 2021 to 2027. Part of the EuroHPC JU is the LUMI supercomputer, which (as of now, 2024) ranks number 5 in the world’s most powerful non-distributed computer systems and number 1 in Europe.

In practice, the users of numerical methods can roughly be grouped into two categories. On the one hand, there are researchers who try to simulate very specific problems and who – out of necessity – develop novel numerical methods to push out the frontier of what can be simulated. On the other hand, there are simulation engineers who use numerical methods (often well established ones, packaged in commercially developed software) as a supporting tool to help make decisions in solving everyday engineering problems. The accuracy of the employed numerical method is typically equally important to both groups. However, efficiency and robustness play a lesser role for researchers in light of developing new methods to simulate problems never approached before. The result is often research-oriented codes developed at universities and research centers. While these codes are typically open-source, they are not necessarily easy to use and often require specific knowledge to operate or even install.

A crucial part of exploiting numerical methods to help develop new products and manufacturing processes or to make existing ones more efficient is the knowledge and technology transfer between research and industry. When viewing simulations as the essential part of the development and production chains that they are, this should be considered a fourth critical aspect next to accuracy, efficiency, and robustness: *accessibility*.

[3]: European Commission (2018), *Council backs Commission’s plans to invest €1 billion in world-class European supercomputers*

1.2 Objectives

This thesis addresses the development and application of advanced numerical methods for simulating metal forming and manufacturing processes. The work aims to enhance the modeling of complex material behaviors, such as isochoric plastic deformation, and the handling of large strains and strain rates. Additionally, the study explores innovative approaches for dealing with complex geometries and moving boundary conditions.

Effective technology transfer between research and industry is essential for translating cutting-edge numerical methods into practical tools for industrial applications. Bridging this gap ensures that innovative

academic advancements can be seamlessly implemented in real-world settings, improving efficiency and accuracy in industrial simulations.

Friction Stir Welding is used as a representative case study, illustrating the relevance and potential of said advanced numerical methods. All methods validated both through numerical convergence studies and by replicating and comparing experimental results.

The individual objectives are:

State-of-the-Art Perform an investigation of the numerical methods used for the simulation of metal forming and manufacturing processes and identify challenges

Finite Element Technology Address isochoric material behavior, observed in the plastic deformation of metals and large strains/strain rates, common to many metal forming processes by employing appropriate finite element technology.

Accuracy and Efficiency Balance numerical accuracy and computational efficiency by employing adaptive refinement schemes targeting both the computational mesh and the finite element formulation itself.

Complex Geometries Study the advantages and challenges of different embedded domain methods to deal with complex and moving boundary conditions and choose a candidate based on its suitability to the mechanical problem at hand and its compatibility with the other, previously introduced, numerical methods.

Industrial Application Apply the introduced methods to Friction Stir Welding as an example of a challenging manufacturing process. Justify the choice of methods by reviewing established approaches to simulating FSW and pointing out a gap between existing methods.

Validation Validate the numerical methods through numerical convergence analysis and comparison to experimental data.

Accessibility Implement the complex numerical methods in a high-performance computing environment to ensure applicability to large-scale industrial problems. Make the generation of input files and interpretation of results accessible enough to be used by industrial simulation engineers rather than academic researchers.

1.3 Outline

The thesis is structured in three major parts. Sec. 2 introduces different Finite Element technologies to deal with isochoric material behavior, stress, and strain accuracy, and a novel adaptive refinement method to enhance efficiency.

Sec. 3 analyses different embedded domain methods and investigates which are most suitable for dealing with the complex geometries and boundary conditions in FSW and other metal forming processes. The previously developed adaptive refinement technique is combined with adaptive mesh refinement to achieve maximum versatility and ensure the highest degrees of accuracy and efficiency.

Sec. 4 introduces Friction Stir Welding (FSW) as a metal joining technology and provides a state-of-the-art review of current simulation approaches, motivating the approach taken in this work. Next a Finite Element framework is developed, designed to address the unique complexities of FSW using an Embedded Domain Method (EDM). Numerical results are presented, followed by experimental validation of the method. Ultimately, the accessibility of the presented framework is addressed, and efforts to facilitate problem specific pre- and post-processing are discussed.

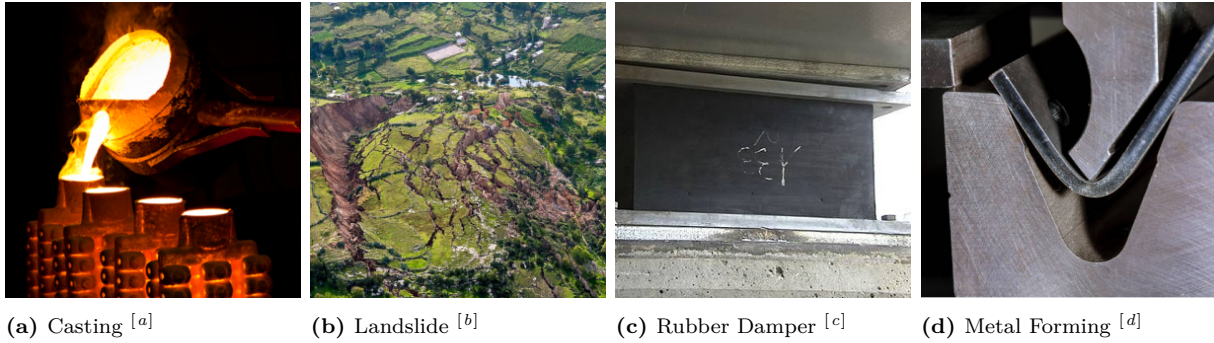
Finally, Sec. 5 highlights the contributions of this thesis, and future lines of research are presented. This covers both opportunities to further FSW simulation techniques and exploit the developed methods for other metal forming and manufacturing applications.

The first part of this work addresses incompressible material behavior as well as strain and stress accuracy and presents existing and novel finite element technologies to overcome challenges, such as numerical instability.

2.1 Objectives

2.1.1 Isochoric Material Behavior

A common occurrence in mechanical problems are isochoric deformations or materials that exhibit isochoric behavior. This means that during a deformation, the volume of a body (solid or fluid) is preserved. In fluid mechanics, this can occur due to the properties of the fluid (Fig. 2.1a) or the flow conditions (Fig. 2.1b).



In solid mechanics, certain materials are known to be *incompressible*, such as bio-tissues or rubber (Fig. 2.1c). They exhibit a Poisson's ratio of $\nu \rightarrow 0.5$, or, in other terms, have much higher bulk modulus than shear modulus $\kappa \gg G$. Another common occurrence of isochoric behavior is the plastic flow of many metals. When deformed elastically, they exhibit compressible behavior with Poisson's ratios typically between 0.2 and 0.35 (see Tab. 2.1). However, once they start to yield, they deform in a volume-preserving manner.

Unfortunately, standard irreducible finite elements do not work in the incompressible domain. Consequently, the behavior of these materials cannot be accurately simulated. Next, one example from linear elasticity and one from J_2 plasticity is presented to highlight the numerical challenges.

Material	Poisson's Ratio ν
Magnesium	0.252 – 0.289
Titanium	0.265 – 0.34
Copper	0.33
Aluminum-alloy	0.32
Stainless Steel	0.3 – 0.31
Cast Iron	0.21 – 0.26

Figure 2.1: Examples of the occurrence of isochoric flow and incompressible materials

^a <https://radheshyaminvestmentcasting.com/...>
(cropped, accessed: 02.05.2025)

^b <https://upload.wikimedia.org/...>
(cropped, accessed: 02.05.2025)

^c <https://bridgebearing.org/...>
(cropped, accessed: 02.05.2025)

^d <https://image.freepik.com/...>
(cropped, accessed: 02.05.2025)

Table 2.1: Poisson's Ratios of common Metals

Incompressible Linear Elasticity The three governing equations of the linear elasticity problem are the Balance of Momentum Eq. (2.1), the Constitutive Eq. (2.2), and the Kinematic Eq. (2.3).

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{b} = \mathbf{0} \quad (2.1)$$

$$\boldsymbol{\sigma} - \mathbb{C} : \boldsymbol{\varepsilon} = \mathbf{0} \quad (2.2)$$

$$\nabla^s \mathbf{u} - \boldsymbol{\varepsilon} = \mathbf{0}, \quad (2.3)$$

where $\boldsymbol{\sigma}$ is the symmetric Cauchy stress tensor, $\mathbf{b}$ is the body force vector, $\mathbb{C}$ is the 4th order stiffness tensor, $\boldsymbol{\varepsilon}$ is the symmetric strain tensor and $\nabla^s \mathbf{u}$ denotes the symmetric gradient of the displacements.

In some situations, it is beneficial to split the stress and strain-related variables into their deviatoric (shape-affecting) and volumetric (volume-affecting) parts such that

$$\boldsymbol{\sigma} = \mathbf{s} + p\mathbf{1}, \quad (2.4)$$

$$\boldsymbol{\varepsilon} = \mathbf{e} + \varepsilon^\circ \mathbf{1}, \quad \text{and} \quad (2.5)$$

$$\mathbb{C} = \mathbb{C}' + \mathbb{C}^\circ. \quad (2.6)$$

The deviatoric parts of the strain and the stress are denoted by $\mathbf{e}$ and $\mathbf{s}$, respectively. The volumetric part of the stress, denoted as p , is also known as the hydrostatic pressure and ε° denotes the volumetric strain. The symbol $\mathbf{1}$ denotes the second order unity tensor δ_{ij} .

In linear elasticity, the stiffness tensor $\mathbb{C}$ can be expressed as a function of either Young's Modulus and Poisson's Ratio $\mathbb{C}(E, \nu)$ or of the shear and bulk modulus $\mathbb{C}(G, \kappa)$. The relationship is as follows.

$$\kappa = \frac{2G(1+\nu)}{3(1-2\nu)} \quad (2.7)$$

For high Poisson Ratios ($0.49 < \nu < 0.5$, i.e., nearly incompressible), the bulk modulus is orders of magnitudes larger than the shear modulus $\kappa \gg G$ and for Poisson's Ratio $\nu = 0.5$ (fully incompressible), the bulk modulus becomes infinity $\kappa = \infty$. The problem becomes apparent when showing the constitutive matrix $\mathbf{C}$ (here in Voigt notation) used in the definition of the element stiffness matrix of the linear elasticity problem

$$\mathbf{K}_h = \int_{\Omega^e} \mathbf{B}^\top \mathbf{C} \mathbf{B} \, dV \quad \text{with} \quad \mathbf{C} = \begin{bmatrix} \kappa + 4G/3 & \kappa - 2G/3 & \kappa - 2G/3 & 0 & 0 & 0 \\ \kappa - 2G/3 & \kappa + 4G/3 & \kappa - 2G/3 & 0 & 0 & 0 \\ \kappa - 2G/3 & \kappa - 2G/3 & \kappa + 4G/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix}. \quad (2.8)$$

As entries of the above matrix become infinite, solving the problem using the standard irreducible finite element method becomes impossible.

The problem with nearly incompressible material behavior is the appearance of volumetric locking (discussed in the next paragraph), which

results in an un-physically stiff response to volumetric deformations and pressure oscillations. Additionally, the linear system becomes extremely ill-conditioned, which degrades the performance of iterative solvers.

J_2 Plasticity Isochoric behavior can also occur during plastic deformation. In the following, this is demonstrated for perfect J_2 plasticity.

It is widely understood that the yielding of metals is governed by the second invariant of the deviatoric stress tensor $J_2 = \frac{1}{2} \mathbf{s} : \mathbf{s}$. Von Mises popularized this approach over a century ago, and it is still used today [4]. In plasticity, an additive decomposition of the infinitesimal strain tensor is considered such that

[4]: Borja (2013), *Plasticity*

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_{\text{el}} + \boldsymbol{\varepsilon}_{\text{pl}}, \quad (2.9)$$

and the elastic constitutive equation is written as

$$\boldsymbol{\sigma} = \mathbb{C}_{\text{el}} : \boldsymbol{\varepsilon}_{\text{el}} = \mathbb{C}_{\text{el}} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_{\text{pl}}). \quad (2.10)$$

The yield surface $f(\boldsymbol{\sigma}, \sigma_Y) = 0$ defines whether a given stress state $\boldsymbol{\sigma}$ causes the material to yield. This depends on the specific yield criterion and a material parameter σ_Y . In von Mises' model, yielding occurs when the second invariant of the deviatoric stress tensor reaches a material-dependent threshold

$$f(\mathbf{s}, \sigma_Y) = \sqrt{2 J_2(\mathbf{s})} - \sigma_Y = 0. \quad (2.11)$$

Hence it is also called J_2 plasticity.

Once the material yields, the plastic flow rule determines the evolution of the plastic strain

$$\dot{\boldsymbol{\varepsilon}}_{\text{pl}} = \dot{\lambda} \frac{\partial f}{\partial \boldsymbol{\sigma}} = \dot{\lambda} \frac{\mathbf{s}}{\|\mathbf{s}\|}, \quad (2.12)$$

where $\mathbf{s}/\|\mathbf{s}\|$ determines the flow direction and $\dot{\lambda}$ determines the magnitude of plastic deformation. It becomes evident from Eq. (2.12) that the evolution of the plastic strain solely depends on the deviatoric stresses, and, therefore, the plastic strain is deviatoric itself.

$$\text{tr}(\dot{\boldsymbol{\varepsilon}}_{\text{pl}}) = \frac{\dot{\lambda}}{\|\mathbf{s}\|} \text{tr}(\mathbf{s}) = 0 \quad (2.13)$$

Typically, plastic strains dominate the total strains, and as a consequence, the elastic strain can be neglected. Unfortunately, finite element interpolation functions cannot properly approximate an isochoric strain field. This is because the displacement field is overconstrained by the additional constraint of preserving the volume throughout the deformation. This effect is commonly known as *volumetric locking*.

Sec. 2.2 introduces different approaches to alleviate the effects of volumetric locking and address fully incompressible materials.

2.1.2 Strain & Stress Accuracy

Strains and stresses are important engineering measurements, and stresses especially play a vital role in predicting plastic deformations and/or failure. In the standard finite element method for solid mechanics problems, displacements are the primary unknown. It is widely known that for a smooth solution, the expected rate of convergence of the global error of the displacements in the L_2 norm is $p + 1$, where p is the polynomial order of the FE interpolation functions (see Tab. 2.2) [5]. However, locally, only a convergence rate of p can be guaranteed. Similarly, the convergence rate of a derivative of the primary unknown, *e.g.*, the strain, is additionally reduced by one.

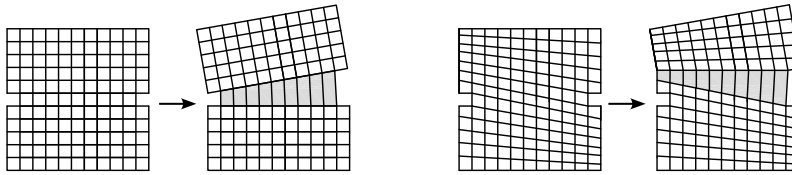
Order p	Global $\mathbf{u}$	Local $\mathbf{u}$	Global $\nabla \cdot \mathbf{u}$	Local $\nabla \cdot \mathbf{u}$
3	4.0	3.0	3.0	2.0
2	3.0	2.0	2.0	1.0
1	2.0	1.0	1.0	—

[5]: Hughes (2000), *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*

Table 2.2: Local and global convergence rates for the primary unknown and a 1st order derivative

Consequently, the local convergence of strains and stresses cannot be guaranteed when using elements of order one, *i.e.*, linear elements.

On the one hand, the approximation of strains and stresses is simply less accurate than the displacements, not least because they become constant within simplicial elements (*i.e.*, triangles, tetrahedra). On the other hand, the non-guaranteed local convergence can have devastating effects when dealing with quasi-singular points such as crack-tips. In fracture mechanics, the stress field determines the direction of crack growth. If local convergence cannot be guaranteed, their evaluation and, ultimately, the crack trajectories can become spuriously mesh dependent, as shown in Fig. 2.2 [6, 7].



[6]: Cervera et al. (2006), “Mesh objective tensile cracking via a local continuum damage model and a crack tracking technique”

[7]: Barbat Vlad (2021), “Mixed finite elements with independent strain interpolation for isotropic and orthotropic damage”

Sec. 2.2 introduces different approaches to improve strain and stress accuracy and ensure local convergence.

Figure 2.2: Mesh dependent crack evolution

Figure origin: [8] (rearranged)

2.2 State Of The Art

In the previous section, three critical limitations of the standard finite element method were identified: coefficient matrix entries tending toward infinity in fully incompressible materials, non-guaranteed local convergence of stresses leading to mesh bias in fracture mechanics, and volumetric locking in cases where the strain field is primarily deviatoric.

Numerous approaches have been developed to address these challenges over recent decades. This section provides an overview of these methods, outlining their contributions, advantages, and limitations.

2.2.1 h , p and hp -refinement

A finer discretization of the computational domain is required to increase the overall solution accuracy. Traditionally, this is achieved by two means. The most straightforward way is to increase the number of elements while keeping the type of element and order of interpolation unchanged. This is known as h -refinement and is depicted in Fig. 2.3b. If the problem is well posed, refining the mesh causes the simulation to converge to the exact solution. However, the finer the mesh, the higher the computational cost. Furthermore, in a previous section (Sec. 2.1.2), it was already established that h -refinement of linear (order 1) elements cannot guarantee the local convergence of strains and stresses. Additionally, it is known that h -refinement can, in specific circumstances, alleviate shear locking but never volumetric locking.

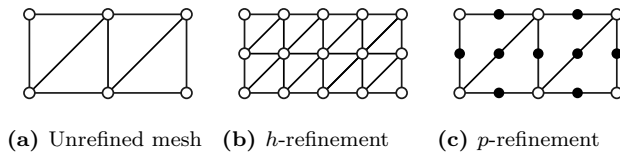


Figure 2.3: Refinement approaches

A different approach is increasing the interpolation function order p . In this case, quadratic, cubic, or even higher order polynomials are used to approximate the solution, which generally leads to more accurate results. As established in Tab. 2.4, quadratic interpolation is already sufficient to ensure local convergence of strains and stresses.

The number of elements may not change, but increasing p also increases computational costs as the number of unknowns and integration points per element rises. Increasing the polynomial order is a good approach if the solution varies smoothly because few coarse, high-order elements are sufficient to capture the solution accurately. However, an excessive increase of the interpolation order can cause effects of over-fitting and oscillations of the solution, especially if the solution is not particularly smooth. Furthermore, the bandwidth of the linear system increases for every additional polynomial order, which is not the case for additional levels of h -refinement. Ultimately, high-order elements can experience a form of hourglassing when subjected to isochoric deformations.

Typically, a combination of h and p -refinement (known as hp -refinement) is used, depending on the problem. Most commercial codes offer quadratic elements, and the rest of the required accuracy is gained through mesh refinement. However, for some specific applications it is very convenient to have linear elements, *e.g.*, when lumping mass matrices, or when employing specialized sub-integration rules. In those cases, the shortcomings of linear elements need to be addressed in other ways.

2.2.2 Reduced and Selective Reduced Integration

Volumetric locking arises from an overconstrained displacement field, but this suggests a straightforward solution: reducing the number of integration points reduces the total number of constraints per element. Zienkiewicz, Taylor, and Too [9] showed that this method could allevi-

[9]: Zienkiewicz et al. (1971), “Reduced integration technique in general analysis of plates and shells”

ate the volumetric locking and, today, this method is known as *reduced integration*. For linear elements, this results in a single integration point located at the center of the element.

$$\underbrace{\int_{\Omega} \nabla^s \delta \mathbf{u} : \mathbb{C} : \nabla^s \mathbf{u} \, dV}_{\text{reduced integration}} = \int_{\Omega} \delta \mathbf{u} \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u} \cdot \mathbf{t} \, dA \quad (2.14)$$

However, reduced integration comes not only at the price of less accurate integration, but it also encourages an effect called *hourglassing*. Since the only remaining integration point is located at the element's center, the resistance against deformation into a trapezoidal shape is greatly reduced. Consequentially, the elements collapse in a particular pattern resembling an hourglass shape, also called a *zero-energy mode*, as the deformation does not contribute to the elements strain energy.

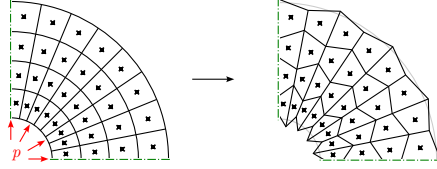


Figure 2.4: Hourglassing in a cylinder under internal pressure due to reduced integration.

Fried [10] and Naylor [11] discovered that only the volumetric part of the total deformation needed to be addressed by reduced integration to overcome the locking problem. By splitting the stiffness contribution of the weak form of the mechanical problem (LHS of Eq. 2.14) into its volumetric and deviatoric part, it is possible to only use reduced integration on the volumetric part, recovering the full integration for the deviatoric part.

$$\underbrace{\int_{\Omega} \nabla^s \delta \mathbf{u} : \mathbb{C}' : \nabla^s \mathbf{u} \, dV}_{\text{deviatoric/full int.}} + \underbrace{\int_{\Omega} \nabla^s \delta \mathbf{u} : \mathbb{C}^o : \nabla^s \mathbf{u} \, dV}_{\text{volumetric/reduced int.}} \quad (2.15)$$

$$= \int_{\Omega} \delta \mathbf{u} \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u} \cdot \mathbf{t} \, dA$$

This approach, known as *selective reduced integration*, leads to an equally locking-free solution with the benefit of a less pronounced hourglassing effect as the fully integrated deviatoric part remains a resistance against the trapezoidal deformation form.

2.2.3 $\bar{\mathbf{B}}$ -Methods

Separating the volumetric and deviatoric parts has proven to be a reliable strategy to address nearly incompressible material behavior. However, the selective reduced integration method relies on splitting the constitutive equation into their respective volumetric and deviatoric parts as shown in Eqs. (2.6) and (2.15). Unfortunately, splitting the constitutive law might not always be possible. This is, for example, the case for anisotropic materials or in some instances of plasticity where the yield criterion and flow rule do not solely depend on the deviatoric stress tensor, as is the case for J_2 plasticity.

[10]: Fried (1974), “Finite element analysis of incompressible material by residual energy balancing”

[11]: Naylor (1974), “Stresses in nearly incompressible materials by finite elements with application to the calculation of excess pore pressures”

The so-called $\bar{\mathbf{B}}$ -Method, which was proposed by Hughes [12], follows the same idea of splitting volumetric and deviatoric parts. However, instead of the constitutive law, the strain definition is split.

[12]: Hughes (1980), “Generalization of selective integration procedures to anisotropic and nonlinear media”

In practical terms, this means splitting the $\mathbf{B}$ -operator,

$$\mathbf{B}_a(\boldsymbol{\xi}) = \begin{bmatrix} B_{a1} & 0 & 0 \\ 0 & B_{a2} & 0 \\ 0 & 0 & B_{a3} \\ B_{a2} & B_{a1} & 0 \\ 0 & B_{a3} & B_{a2} \\ B_{a3} & 0 & B_{a1} \end{bmatrix} \quad \text{with} \quad B_{ai} = \frac{\partial N_a(\boldsymbol{\xi})}{\partial x_i} \quad \text{at Node } a, \quad (2.16)$$

which is responsible for mapping nodal displacements $\mathbf{u}_a$ to strains in Voigt notation $\boldsymbol{\varepsilon}_V$ at the integration points. $\boldsymbol{\xi}$ denotes the natural coordinates.

$$\boldsymbol{\varepsilon}_V(\boldsymbol{\xi}) = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{23} \\ \varepsilon_{13} \end{bmatrix} = \mathbf{B}_a(\boldsymbol{\xi}) \cdot \mathbf{u}_a \quad (2.17)$$

The volumetric-deviatoric split can be shifted from the constitutive law to the strain definition because volumetric and deviatoric parts are orthogonal to each other. Building on Eqs. (2.4) to (2.6), it can be shown that

$$\mathbf{s} = \mathbb{C}' : \boldsymbol{\varepsilon} = \mathbb{C} : \mathbf{e} \quad \text{and} \quad \mathbf{0} = \mathbb{C}' : \varepsilon^\circ \mathbf{1} = \mathbb{C}^\circ : \mathbf{e}. \quad (2.18)$$

Consequently, the $\mathbf{B}$ -operator can be split as follows

$$\mathbf{B}_a = \mathbf{B}_a^{\text{vol}} + \mathbf{B}_a^{\text{dev}} \quad (2.19)$$

$$\mathbf{B}_a^{\text{vol}} = \frac{1}{3} \begin{bmatrix} B_{a1} & B_{a2} & B_{a3} \\ B_{a1} & B_{a2} & B_{a3} \\ B_{a1} & B_{a2} & B_{a3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B}_a^{\text{dev}} = \frac{1}{3} \begin{bmatrix} 2B_{a1} & -B_{a2} & -B_{a3} \\ -B_{a1} & 2B_{a2} & -B_{a3} \\ -B_{a1} & -B_{a2} & 2B_{a3} \\ 3B_{a2} & 3B_{a1} & 0 \\ 0 & 3B_{a3} & 3B_{a2} \\ 3B_{a3} & 0 & 3B_{a1} \end{bmatrix} \quad (2.20)$$

Ultimately, the volumetric part $\mathbf{B}_a^{\text{vol}}$ is modified to alleviate locking.

This can be achieved in several ways, which is why the $\bar{\mathbf{B}}$ -Method is known as a generalization of other methods. Depending on the modification applied to $\mathbf{B}_a^{\text{vol}}$, different approaches are replicated. By evaluating the shape functions at the center of the element

$$\bar{B}_{ai}(\boldsymbol{\xi}) = B_{ai}(\mathbf{0}) = \left. \frac{\partial N_a(\boldsymbol{\xi})}{\partial x_i} \right|_{\boldsymbol{\xi}=\mathbf{0}}, \quad (2.21)$$

the selective reduced integration is recovered.

An alternative approach replicates what is known as the *Mean Dilata-*

tion Formulation. Here, the value of B_{ai} , evaluated at each integration point, is replaced by its element average

$$\bar{B}_{ai} = \int_{\Omega^e} B_{ai} \, dV \cdot \left(\int_{\Omega^e} dV \right)^{-1} \quad (2.22)$$

The weak form of the small strain solid mechanics problem can be expressed as follows

$$\begin{aligned} \int_{\Omega} \delta \mathbf{u}_h (\mathbf{B}^{\text{dev}} + \bar{\mathbf{B}}^{\text{vol}})^{\top} : \mathbb{C} : (\mathbf{B}^{\text{dev}} + \bar{\mathbf{B}}^{\text{vol}}) \mathbf{u}_h \, dV \\ = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \mathbf{t} \, dA, \end{aligned} \quad (2.23)$$

which, crucially, leaves the constitutive tensor $\mathbb{C}$ unmodified.

$\bar{\mathbf{B}}$ -Methods have been generalized for hyperelastic finite-deformation formulations by Hughes, Taylor, and Sackman [13] and Simo, Taylor, and Pister [14], where the target of the volumetric split is the deformation gradient $\mathbf{F}$; hence they are known as $\bar{\mathbf{F}}$ -Methods.

2.2.4 Enhanced Assumed Strain (EAS)

As established before, the underlying problem of locking is overconstraining available degrees of freedom. While methods based on reduced integration address this problem by reducing the number of integration points, which consequentially reduces the total number of constraints per element, the *enhanced assumed strains* (EAS) method takes a different approach. The EAS method artificially introduces additional DOFs $\boldsymbol{\alpha}$ to the element, which represent different deformation modes of the element linked to different strain configurations. A linear 3D hexahedral element with eight nodes and 24 total degrees of freedom can amount to up to 30 additional EAS modes [15]. Each mode adds additional flexibility to the element, overcoming the overconstraining.

Simo and Rifai [16] proposed to amend the classical, compatible strains by an additional strain such that

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^{\text{com}} + \tilde{\boldsymbol{\varepsilon}}. \quad (2.24)$$

The classical compatible strain is computed as usual from the displacements using the symmetric gradient or, in discretized form, the $\mathbf{B}$ -operator.

$$\boldsymbol{\varepsilon}^{\text{com}} = \nabla^s \mathbf{u} = \mathbf{B} \cdot \mathbf{U}, \quad (2.25)$$

where $\mathbf{U}$ is the vector of the nodal unknowns. The enhanced assumed strains are computed in a similar fashion

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{G} \cdot \boldsymbol{\alpha} = \frac{J_0}{J(\boldsymbol{\xi})} \cdot \mathbf{F}_0^{-\top} \cdot \mathbf{M}(\boldsymbol{\xi}) \cdot \boldsymbol{\alpha}, \quad (2.26)$$

where J is the Jacobian determinant of the transformation matrix between spacial and intrinsic coordinates, J_0 is said determinant, evaluated at the element center, $\mathbf{F}_0$ is a 6×6 spacial-intrinsic-coordinate-

[13]: Hughes et al. (1974), *Finite Element Formulation and Solution of Contact-Impact Problems in Continuum Mechanics*

[14]: Simo et al. (1985), “Variational and projection methods for the volume constraint in finite deformation elasto-plasticity”

[15]: Duchêne et al. (2011), “Assessment of the Enhanced Assumed Strain (EAS) and the Assumed Natural Strain (ANS) techniques in the mechanical behavior of the SSH3D Solid-Shell element”

[16]: Simo et al. (1990), “A class of mixed assumed strain methods and the method of incompatible modes”

transformation matrix for strains [17] and, finally, $\mathbf{M}(\boldsymbol{\xi})$ a matrix containing potential EAS modes for each of the six independent strains.

In practice, it is up to the user to decide which modes to consider. Note that the actual number of unknowns does not change since the additional DOFs (α values) are eliminated at the element level by means of static condensation before assembly.

2.2.5 Q_1P_0

The overarching theme of methods to address isochoric behavior is splitting the volumetric and deviatoric parts. In the Selective Reduced Integration approach, the split is applied to the constitutive law, whereas in the $\bar{\mathbf{B}}$ -method, the strain operator is split. Both methods then treat the volumetric component, which causes the locking effects in nearly incompressible materials, differently from the deviatoric component.

The following method takes this approach further and splits all governing equations. This results, among other things, in the splitting of the stress into its deviatoric part and the pressure, see Eq. (2.4). The novelty of this method is that it treats the pressure as a primary unknown next to displacements. The complete derivation is omitted at this point and shown instead in Sec. 2.3, for reasons that become clear shortly.

The displacements are then interpolated using linear shape functions, whereas the pressure is assumed constant within the element. This type of element is commonly known as Q_1P_0 (for quadrilaterals) or P_1P_0 (for triangles). This element is the first representative of a class of *mixed finite elements*, which refers to having more than one primary unknown (remember, the additional unknowns of the EAS method were eliminated through static condensation). However, the Q_1P_0 element is prone to instabilities for fully isochoric materials, a problem it shares with many mixed formulations, as will be addressed in Sec. 2.4.

The reason the Q_1P_0 element was introduced separately (as opposed to as part of the general mixed two-field formulation) is the fact that, for nearly incompressible problems, the element is equivalent to Selective Reduced Integration, as was shown by Hughes [18].

2.3 Mixed Formulations

The volumetric-deviatoric split of the governing equations is the starting point for developing a mixed finite element formulation, free of volumetric locking and capable of dealing with fully isochoric material behavior.

The volumetric and deviatoric parts of the strain, stress, and constitutive tensors can be easily extracted using the 4th-order volumetric and deviatoric operators

$$\mathbb{V} = \frac{1}{3} \mathbf{I} \otimes \mathbf{I}, \quad (2.27)$$

$$\mathbb{D} = \mathbb{I} - \mathbb{V}, \quad (2.28)$$

[17]: Alves de Sousa et al. (2007), “On the use of a reduced enhanced solid-shell (RESS) element for sheet forming simulations”

[18]: Hughes (1977), “Equivalence of Finite Elements for Nearly Incompressible Elasticity”

where $\mathbb{I}$ is the 4th-order identity $[\delta_{ik}\delta_{jl}]$ and $\mathbf{I}$ is the 2nd-order identity $[\delta_{ij}]$, with δ_{ij} being Kronecker's delta. Note that $\mathbf{I} \otimes \mathbf{I}$ can be expressed as $[\delta_{ij}\delta_{kl}] \neq \mathbb{I}$ and is *not* an identity.

Using these operators, it is possible to extract the volumetric and deviatoric parts of 2nd and 4th-order tensors such that

$$\begin{aligned}\boldsymbol{\sigma} &= \mathbb{V} : \boldsymbol{\sigma} + \mathbb{D} : \boldsymbol{\sigma} \\ &= p\mathbf{I} + \mathbf{s}\end{aligned}\quad (2.29)$$

and

$$\begin{aligned}\boldsymbol{\varepsilon} &= \mathbb{V} : \boldsymbol{\varepsilon} + \mathbb{D} : \boldsymbol{\varepsilon} \\ &= \frac{1}{3}\boldsymbol{\varepsilon}^\circ \mathbf{I} + \mathbf{e}.\end{aligned}\quad (2.30)$$

Finally, the constitutive tensor can be split in an equal fashion

$$\begin{aligned}\mathbb{C} &= \mathbb{V} : \mathbb{C} + \mathbb{D} : \mathbb{C} \\ &= \mathbb{C}^\circ + \mathbb{C}'.\end{aligned}\quad (2.31)$$

Note that in linear elasticity the volumetric and deviatoric constitutive tensors can be expressed in terms of their respective operators and associated moduli (bulk and shear)

$$\mathbb{C} = 3\kappa \mathbb{V} + 2G \mathbb{D}. \quad (2.32)$$

After a series of operations, the governing Eqs. (2.1) to (2.3) can now be expressed as

$$\nabla \cdot \mathbf{s} + \nabla p + \mathbf{b} = 0 \quad (2.33)$$

$$\mathbf{s} - \mathbb{C}' : \mathbf{e} = \mathbf{0} \quad (2.34)$$

$$\mathbb{D} : \nabla^s \mathbf{u} - \mathbf{e} = \mathbf{0} \quad (2.35)$$

$$\nabla \cdot \mathbf{u} - p/\kappa = 0, \quad (2.36)$$

where Eq. (2.33) is the balance of momentum equation in mixed form, Eq. (2.34) is the deviatoric part of the constitutive equation, Eq. (2.35) is the deviatoric part of the kinematic equation and Eq. (2.36) is the volumetric part of the kinematic equation where the term p/κ originates from the volumetric part of the constitutive equation.

In the incompressible limit ($\kappa = \infty$), Eq. (2.36) is known as the *kinematic* or *incompressibility constraint* and enforces volume preservation ($\nabla \cdot \mathbf{u} = 0$). Furthermore, the constitutive tensor no longer contains infinite values in its coefficient matrix, because it no longer depends on κ . In Voigt notation, it can be expressed as

$$\mathbf{C}' = \begin{bmatrix} 4G/3 & -2G/3 & -2G/3 & 0 & 0 & 0 \\ -2G/3 & 4G/3 & -2G/3 & 0 & 0 & 0 \\ -2G/3 & -2G/3 & 4G/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix}. \quad (2.37)$$

2.3.1 Two-field u/p Formulation

Derivation By introducing Eqs. (2.34) and (2.35) into the balance of momentum Eq. (2.33), the strong form of the mixed u/p problem is obtained. It consists of finding the unknowns $\mathbf{u}$ and p to satisfy

$$\nabla \cdot (\mathbb{C}' : \nabla^s \mathbf{u}) + \nabla p + \mathbf{b} = \mathbf{0} \quad (2.38)$$

$$\nabla \cdot \mathbf{u} - p/\kappa = 0. \quad (2.39)$$

The problem is complemented by the boundary conditions

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{on } \Gamma_D, \quad (2.40)$$

$$\mathbf{s} \cdot \mathbf{n} + p \mathbf{n} = \bar{\mathbf{t}} \quad \text{on } \Gamma_N, \quad (2.41)$$

where $\mathbf{n}$ is the outward normal vector of the boundary $\Gamma = \Gamma_D \cup \Gamma_N$. Dirichlet and Neumann boundaries do not intersect $\Gamma_D \cap \Gamma_N = \emptyset$.

The weak form of the mixed u/p problem is obtained after multiplication with the corresponding test functions $\delta \mathbf{u}$ and δp and integration by parts. It reads

$$\begin{aligned} \int_{\Omega} \nabla^s \delta \mathbf{u} : \mathbb{C}' : \nabla^s \mathbf{u} \, dV + \int_{\Omega} (\nabla \cdot \delta \mathbf{u}) \, p \, dV \\ + \int_{\Omega} \delta p (\nabla \cdot \mathbf{u}) \, dV - \int_{\Omega} \delta p \, p / \kappa \, dV \\ = \int_{\Omega} \delta \mathbf{u} \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u} \cdot \bar{\mathbf{t}} \, dA, \end{aligned} \quad (2.42)$$

The Galerkin discrete version of the above problem is obtained by introducing the approximated unknowns $\mathbf{u}_h = \mathbf{N}_u \mathbf{U}_h$ and $p_h = \mathbf{N}_p \mathbf{P}_h$, where $\mathbf{U}_h$ and $\mathbf{P}_h$ are the nodal values of the velocity and pressure fields. The test functions are defined using the same set of shape functions (*i.e.*, Galerkin's Method), hence $\delta \mathbf{u}_h = \mathbf{N}_u \delta \mathbf{U}_h$ and $\delta p_h = \mathbf{N}_p \delta \mathbf{P}_h$. Note that the equal-order polynomial shape functions are assumed for all unknowns. Hence, from hereon, the index of $\mathbf{N}$ is omitted. The discretized problem reads

$$\begin{aligned} \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV + \int_{\Omega} (\nabla \cdot \delta \mathbf{u}_h) \, p_h \, dV \\ + \int_{\Omega} \delta p_h (\nabla \cdot \mathbf{u}_h) \, dV - \int_{\Omega} \delta p_h \, p_h / \kappa \, dV \\ = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA. \end{aligned} \quad (2.43)$$

The elemental stiffness matrix and corresponding right-hand side read

$$\mathbf{K}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{B}^\top \mathbb{C}' \mathbf{B} & \mathbf{G} \mathbf{N} \\ \mathbf{N}^\top \mathbf{G}^\top & -\mathbf{N}^\top \mathbf{N} / \kappa \end{bmatrix} \quad \text{and} \quad (2.44)$$

$$\mathbf{F}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{N}^\top \mathbf{b} + \mathbf{N}_{\Gamma_N}^\top \bar{\mathbf{t}} \\ \mathbf{0} \end{bmatrix}, \quad (2.45)$$

where $\mathbf{B}$ and $\mathbf{B}^\top$ are the symmetric gradient and divergence operators

and $\mathbf{G}$ and $\mathbf{G}^\top$ are the (regular) gradient and divergence operators, respectively. $\mathbf{N}_{\Gamma_N}^\top$ signals integration over the element's boundary.

Unfortunately, the mixed u/p formulation is not naturally stable in the incompressible limit for arbitrary orders of polynomial shape functions. This includes linear/linear (Q_1Q_1/P_1P_1) elements, which would be a convenient choice for their simplicity, as well as the aforementioned linear/constant (Q_1P_0/P_1P_0) elements (see Sec. 2.2.5). Sec. 2.4 addresses the reason for this lack of stability and methods to restore it.

Assuming the stability problem is dealt with, it becomes clear why the u/p formulation is a powerful tool for addressing isochoric problems. By adding an additional primary unknown, a higher number of degrees of freedom enhances the “flexibility” of the element, which alleviates the volumetric locking effects of nearly incompressible problems. At the same time, by choosing the pressure as this additional unknown (rather than, *e.g.*, a strain measure), the bulk modulus appears in the denominator of Eqs. (2.39) and (2.42–2.45). Hence, in the incompressible limit, this term disappears.

Equivalences in Fluid Mechanics Incompressible fluid mechanics, typically modeled using the Navier-Stokes Eqs. (2.46) and (2.47), faces the exact same problem. In fact, the approach to treat the pressure as an independent variable (and stabilizing the resulting instabilities) originated from fluid mechanics and is only “recently” (that is, since the 2000s) used in solid mechanics.

The Navier-Stokes equations read

$$-\frac{\partial \dot{\mathbf{u}}}{\partial t} - \dot{\mathbf{u}} \cdot \nabla \dot{\mathbf{u}} + \nabla \cdot \mathbf{s} + \nabla p + \mathbf{b} = \mathbf{0} \quad (2.46)$$

$$\nabla \cdot \dot{\mathbf{u}} = 0. \quad (2.47)$$

with the constitutive law

$$\mathbf{s} = 2\mu \nabla^s \dot{\mathbf{u}}. \quad (2.48)$$

The balance of momentum equation (2.46) shares the reaction term ($\nabla \cdot \mathbf{s} + \nabla p$) with the solid mechanics problem in Eq. (2.49). The counterpart of the kinematic or incompressibility constraint of the solid mechanics problem is often referred to as the *continuity* equation (2.47) in incompressible fluid mechanics.

Remark 2.3.1. *Note that the same convention for splitting the stress in deviatoric and volumetric parts has been maintained as for the solid mechanics problem: $\boldsymbol{\sigma} = \mathbf{s} + p\mathbf{1}$. Typically the pressure is introduced with the opposite sign in fluid mechanics, however, to keep the introduced expressions consistent the same convention as in solid mechanics is kept.*

When omitting the transient ($\partial \dot{\mathbf{u}}/\partial t$) and convective terms ($\dot{\mathbf{u}} \cdot \nabla \dot{\mathbf{u}}$), so-called Stokes flow is obtained. The temporal derivative is typically omitted when looking for a steady state solution and the convective

term when dealing with flow conditions where the viscous forces dominate, *i.e.*, low Reynolds numbers occur. The Stokes flow problem, therefore, reads

$$\nabla \cdot (2\mu \nabla^s \dot{\mathbf{u}}) + \nabla p + \mathbf{b} = \mathbf{0} \quad (2.49)$$

$$\nabla \cdot \dot{\mathbf{u}} = 0. \quad (2.50)$$

From a numerical point of view, this problem is equivalent to incompressible small strain elasticity, with only a small difference. The constitutive law, Eq. (2.48), relates strain rates and stresses with a scalar (the dynamic viscosity μ) rather than the 4th order tensor in incompressible solid mechanics that relates strains and stresses (compare to Eq. (2.34)).

The derivation of the Galerkin discrete version of the stated problem is straightforward.

$$\begin{aligned} \int_{\Omega} \nabla^s \delta \dot{\mathbf{u}}_h : 2\mu \nabla^s \dot{\mathbf{u}}_h \, dV + \int_{\Omega} \nabla \cdot \delta \dot{\mathbf{u}}_h p_h \, dV \\ + \int_{\Omega} \delta p_h \nabla \cdot \dot{\mathbf{u}}_h \, dV \\ = \int_{\Omega} \delta \dot{\mathbf{u}}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \dot{\mathbf{u}}_h \cdot \bar{\mathbf{t}} \, dA, \end{aligned} \quad (2.51)$$

The elemental stiffness matrix and corresponding right-hand side read

$$\mathbf{K}_h = \int_{\Omega^e} \begin{bmatrix} 2\mu \mathbf{B}^T \mathbf{B} & \mathbf{G} \mathbf{N} \\ \mathbf{N}^T \mathbf{G}^T & \mathbf{0} \end{bmatrix} \quad \text{and} \quad (2.52)$$

$$\mathbf{F}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{N}^T \mathbf{b} + \mathbf{N}_{\Gamma_N}^T \bar{\mathbf{t}} \\ \mathbf{0} \end{bmatrix}. \quad (2.53)$$

Application The mixed and stabilized displacement/pressure (u/p) formulation has proven to be a useful tool for simulating isochoric problems. Examples include simple incompressibility in linear elasticity [19], poroelastic materials [20], incompressible plasticity [21], specifically shear-band localization in J_2 plasticity [1, 22], studying of structural size effect in J_2 plasticity [23], in non-linear solid mechanics, including shear-localization [24, 25], topology optimization in the incompressible domain [26–28], in conjunction with phase-field models for finite strain fracture of nearly incompressible hyperelastic materials [29], hyperelasticity [30, 31], friction stir welding [32, 33] and multi-material problems with compressible and incompressible material models [34].

2.3.2 Three-field $u/p/e$ Formulation

Derivation The starting point of developing the mixed u/p formulation was the introduction of Eqs. (2.34) and (2.35) into the balance of momentum Eq. (2.33). By only introducing the constitutive Eq. (2.34) into the balance of momentum equation and keeping the kinematic equation separate, the strong form of the three-field $u/p/e$ problem is obtained [35]. It consists of finding the unknowns $\mathbf{u}$, p and $\mathbf{e}$ (deviatoric

[19]: Chiumenti et al. (2002), “A stabilized formulation for incompressible elasticity using linear displacement and pressure interpolations”

[20]: Atalla et al. (1998), “A mixed displacement-pressure formulation for poroelastic materials”

[21]: Cervera et al. (2003), “Mixed linear/linear simplicial elements for incompressible elasticity and plasticity”

[1]: Cervera et al. (2004), “Shear band localization via local J_2 continuum damage mechanics”

[22]: Cervera et al. (2004), “Softening, localization and stabilization: capture of discontinuous solutions in J_2 plasticity”

[23]: Cervera et al. (2009), “Size effect and localization in J_2 plasticity”

[24]: Cervera et al. (2010), “Mixed stabilized finite element methods in nonlinear solid mechanics. Part I: Formulation”

[25]: Cervera et al. (2010), “Mixed stabilized finite element methods in nonlinear solid mechanics. Part II: Strain localization”

[26]: Castañar et al. (2020), “A stabilized mixed finite element approximation for incompressible finite strain solid dynamics using a total Lagrangian formulation”

[27]: Castañar et al. (2022), “Topological derivative-based topology optimization of incompressible structures using mixed formulations”

[28]: Castañar et al. (2024), “Topology optimization of incompressible structures subject to fluid–structure interaction”

[29]: Tian et al. (2022), “Mixed displacement–pressure–phase field framework for finite strain fracture of nearly incompressible hyperelastic materials”

[30]: Kadapa et al. (2020), “A linearized consistent mixed displacement–pressure formulation for hyperelasticity”

[31]: Codina et al. (2024), “Finite element approximation of stabilized mixed models in finite strain hyperelasticity involving displacements and stresses and/or pressure—An overview of alternatives”

[32]: Chiumenti et al. (2013), “Numerical modeling of friction stir welding processes”

[33]: Saracibar et al. (2014), “Computational Modeling and Sub-Grid Scale Stabilization of Incompressibility and Convection in the Numerical Simulation of Friction Stir Welding Processes”

strains) to satisfy

$$\nabla \cdot (\mathbb{C}' : \mathbf{e}) + \nabla p + \mathbf{b} = 0 \quad (2.54)$$

$$\mathbb{D} : \nabla^s \mathbf{u} - \mathbf{e} = \mathbf{0} \quad (2.55)$$

$$\nabla \cdot \mathbf{u} - p/\kappa = 0. \quad (2.56)$$

The weak form of the mixed $u/p/e$ problem is obtained after multiplication with the corresponding test functions $\delta \mathbf{u}$, δp and $\delta \mathbf{e}$ and integration by parts. It reads

$$\begin{aligned} & \int_{\Omega} (\nabla \cdot \delta \mathbf{u}) p \, dV + \int_{\Omega} \nabla^s \delta \mathbf{u} : \mathbb{C}' : \mathbf{e} \, dV \\ & + \int_{\Omega} \delta p (\nabla \cdot \mathbf{u}) \, dV - \int_{\Omega} \delta p p / \kappa \, dV \\ & + \int_{\Omega} \delta \mathbf{e} : \mathbb{C}' : \nabla^s \mathbf{u} \, dV - \int_{\Omega} \delta \mathbf{e} : \mathbb{C} : \mathbf{e} \, dV \\ & = \int_{\Omega} \delta \mathbf{u} \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u} \cdot \bar{\mathbf{t}} \, dA, \end{aligned} \quad (2.57)$$

The Galerkin discrete version of the above problem is obtained by introducing the approximated unknowns $\mathbf{u}_h = \mathbf{N}_u \mathbf{U}$, $p_h = \mathbf{N}_p \mathbf{P}_h$ and $\mathbf{e}_h = \mathbf{N}_e \mathbf{E}_h$, where $\mathbf{U}_h$, $\mathbf{P}_h$ and $\mathbf{E}_h$ are the nodal values of the displacement, pressure and strain fields. The test functions are defined using the same set of shape functions (*i.e.*, Galerkin's Method), hence $\delta \mathbf{u}_h = \mathbf{N}_u \delta \mathbf{U}_h$, $\delta p_h = \mathbf{N}_p \delta \mathbf{P}_h$ and $\delta \mathbf{e}_h = \mathbf{N}_e \delta \mathbf{E}_h$. Note that the equal-order polynomial shape functions are assumed for all unknowns. Hence, the index of $\mathbf{N}$ is omitted from hereon. The discretized problem reads

$$\begin{aligned} & \int_{\Omega} (\nabla \cdot \delta \mathbf{u}_h) p_h \, dV + \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \mathbf{e}_h \, dV \\ & + \int_{\Omega} \delta p_h (\nabla \cdot \mathbf{u}_h) \, dV - \int_{\Omega} \delta p_h p_h / \kappa \, dV \\ & + \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV - \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C} : \mathbf{e}_h \, dV \\ & = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA. \end{aligned} \quad (2.58)$$

The elemental stiffness matrix and corresponding right-hand side read

$$\mathbf{K}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{0} & \mathbf{GN} & \mathbf{B}^\top \mathbf{C}' \mathbf{N} \\ \mathbf{N}^\top \mathbf{G}^\top & -\mathbf{N}^\top \mathbf{N} / \kappa & \mathbf{0} \\ \mathbf{N}^\top \mathbf{C}' \mathbf{B} & \mathbf{0} & -\mathbf{N}^\top \mathbf{CB} \end{bmatrix} \quad \text{and} \quad (2.59)$$

$$\mathbf{F}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{N}^\top \mathbf{b} + \mathbf{N}_{\Gamma_N}^\top \bar{\mathbf{t}} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (2.60)$$

Similar to the u/p problem, the three-field problem is not stable for arbitrary orders of shape functions. Sec. 2.4 explains the source of this instability and methods to stabilize the formulation. The benefits of the $u/p/e$ formulation are three-fold:

1. By treating the deviatoric strain as a primary unknown, the *accu-*

[34]: Kadapa (2024), “Mixed displacement–pressure formulations and suitable finite elements for multi-material problems with compressible and incompressible models”

[35]: Chiumenti et al. (2021), “Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation”

racy of its approximation is increased, as it is no longer computed by postprocessing the displacements. Therefore, the same is true for the deviatoric stresses, as they are now computed from more accurate strains. This becomes important for problems where not only high strains occur but they are also localized (*e.g.*, softening or shear-thinning).

2. Closely related is the fact that the *convergence rate* of the deviatoric strain is increased. This means that even for linear elements, it is now possible to ensure local convergence of stresses and strains, which is essential for eliminating the mesh bias in cracking problems. The actual obtained rate of convergence can vary depending on the stabilization (see Sec. 2.4), but with the right approach, a convergence rate larger than one is achievable.
3. The ability to tackle fully isochoric problems is preserved due to the pressure being a primary unknown.

Equivalences in Fluid Mechanics The same approach can be used to solve the Navier-Stokes equations, as well as the Stokes problem [35]. The derivation of the discrete Galerkin version of the three-field Stokes flow problem is straightforward and reads

$$\begin{aligned}
 & \int_{\Omega} (\nabla \cdot \delta \dot{\mathbf{u}}_h) p_h \, dV + \int_{\Omega} \nabla^s \delta \dot{\mathbf{u}}_h : 2\mu \dot{\mathbf{e}}_h \, dV \\
 & \quad + \int_{\Omega} \delta p_h (\nabla \cdot \dot{\mathbf{u}}_h) \, dV \\
 & + \int_{\Omega} \delta \dot{\mathbf{e}}_h : 2\mu \nabla^s \dot{\mathbf{u}}_h \, dV - \int_{\Omega} \delta \dot{\mathbf{e}}_h : 2\mu \dot{\mathbf{e}}_h \, dV \\
 & \quad = \int_{\Omega} \delta \dot{\mathbf{u}}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \dot{\mathbf{u}}_h \cdot \bar{\mathbf{t}} \, dA.
 \end{aligned} \tag{2.61}$$

The elemental stiffness matrix and corresponding right-hand side read

$$\mathbf{K}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{0} & \mathbf{GN} & 2\mu \mathbf{B}^\top \mathbf{N} \\ \mathbf{N}^\top \mathbf{G}^\top & \mathbf{0} & \mathbf{0} \\ 2\mu \mathbf{N}^\top \mathbf{B} & \mathbf{0} & -2\mu \mathbf{N}^\top \mathbf{B} \end{bmatrix} \quad \text{and} \tag{2.62}$$

$$\mathbf{F}_h = \int_{\Omega^e} \begin{bmatrix} \mathbf{N}^\top \mathbf{b} + \mathbf{N}_{\Gamma_N}^\top \bar{\mathbf{t}} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{2.63}$$

Application Including a strain-based measure as a primary unknown has proven useful in addressing strain and stress accuracy, especially for linear elements. However, the $u/p/e$ formulation is only one example, next to u/ε (displacement-strain) [24], u/σ (displacement-stress) [24, 36] and $u/p/s$ (displacement-pressure-deviatoric stress) [37]. These formulations all serve a similar purpose; however, typically, the (deviatoric) strain is chosen over the (deviatoric) stress since most constitutive laws are strain-driven. As mentioned before, including the pressure allows approaching the incompressible limit but requires the constitutive law to be split into volumetric and deviatoric parts. Hence for non- J_2 plasticity applications the u/ε formulation is the obvious choice.

[35]: Chiumenti et al. (2021), “Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation”

[24]: Cervera et al. (2010), “Mixed stabilized finite element methods in nonlinear solid mechanics. Part I: Formulation”

[24]: Cervera et al. (2010), “Mixed stabilized finite element methods in nonlinear solid mechanics. Part I: Formulation”

[36]: Scovazzi et al. (2017), “A velocity/stress mixed stabilized nodal finite element for elastodynamics: Analysis and computations with strongly and weakly enforced boundary conditions”

[37]: Chiumenti et al. (2015), “A mixed three-field FE formulation for stress accurate analysis including the incompressible limit”

Mixed and stabilized formulations which include a strain or stress measure, have been used in the past to address various problems. A typical application field for the mixed u/ε formulation are cracking problems including seismic responses to masonry structures [38], multi-cracking [39] and the studying of the size effect [40] and the $u/p/e$ formulation has been used to study structural failure under incompressible conditions [41]. Trinh et al. [42] use the mixed u/ε formulation to predict the flexural behavior of laminated composite beams. Aguirre, Codina, and Baiges [43] have applied the u/σ formulation to solid-shell elements to study thin structures and have extended the analysis to hyperelasticity [44]. Codina, Castañar, and Baiges [31] have used the u/σ and $u/p/s$ formulations in the context of finite strain hyperelasticity and Moreno et al. [45] and Moreno, Codina, and Baiges [46] have used the $u/p/s$ formulation for viscoelastic flow and Itou, Kovtunencko, and Rudoy [47] to predict non-penetrating cracks.

[38]: Vlachakis et al. (2019), “Out-of-plane seismic response and failure mechanism of masonry structures using finite elements with enhanced strain accuracy”

[39]: Cervera et al. (2020), “Architecture of a multi-crack model with full closing, reopening and sliding capabilities”

[40]: Barbat et al. (2020), “Structural size effect: Experimental, theoretical and accurate computational assessment”

[41]: Moreira et al. (2022), “Accurate thermal-induced structural failure analysis under incompressible conditions”

[42]: Trinh et al. (2020), “A strain-displacement mixed formulation based on the modified couple stress theory for the flexural behaviour of laminated beams”

[43]: Aguirre et al. (2024), “Stress-displacement stabilized finite element analysis of thin structures using solid-shell elements, Part I: On the need of interpolating the stresses”

[44]: Aguirre et al. (2024), “Stress-displacement stabilized finite element analysis of thin structures using Solid-Shell elements, Part II: Finite strain hyperelasticity”

[31]: Codina et al. (2024), “Finite element approximation of stabilized mixed models in finite strain hyperelasticity involving displacements and stresses and/or pressure—An overview of alternatives”

[45]: Moreno et al. (2019), “Logarithmic conformation reformulation in viscoelastic flow problems approximated by a VMS-type stabilized finite element formulation”

[46]: Moreno et al. (2020), “Solution of transient viscoelastic flow problems approximated by a term-by-term VMS stabilized finite element formulation using time-dependent subgrid-scales”

[47]: Itou et al. (2021), “Three-field mixed formulation of elasticity model nonlinear in the mean normal stress for the problem of non-penetrating cracks in bodies”

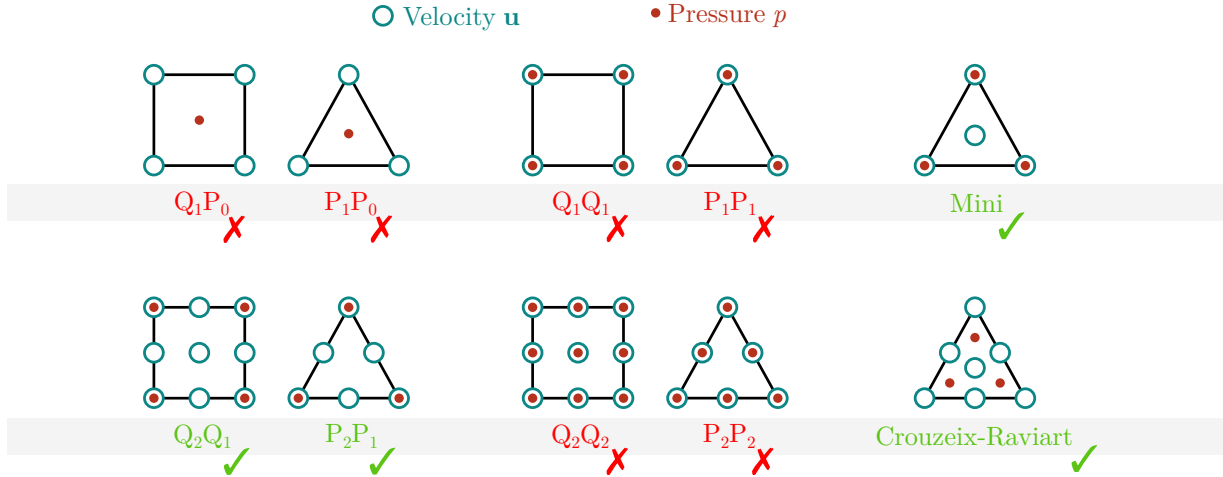
2.4 Numerical Stability and Stabilization

Naturally Stable Elements Unfortunately, mixed finite element formulations are only naturally stable (*i.e.*, they satisfy the *inf-sup* condition) for a particular combination of orders of polynomial shape functions. This condition was first formulated by Brezzi and Fortin [48] and is also known as the Ladyzhenskaya-Babuška-Brezzi (LBB) condition. From a purely mathematical point of view it is a sufficient condition for a saddle point problem (such as linear elasticity or Stokes flow in mixed form) to have a unique solution and that the resulting problem is stable.

As mentioned in previous sections, the linear/linear and linear/constant displacement-pressure u/p formulations do not satisfy the *inf-sup* condition. A commonly used alternative is the quadratic/linear (Q_2Q_1 or P_2P_1) element (also known as the Taylor-Hood element) [49], as it is naturally stable. Fig. 2.5 gives an overview of stable and unstable elements.

[48]: Brezzi et al. (1991), *Mixed and Hybrid Finite Element Methods*

[49]: Taylor et al. (1973), “A numerical solution of the Navier-Stokes equations using the finite element technique”



In the past, researchers have developed elements that were specifically engineered to be naturally stable [50]. One example is the so-called *Mini* element by Arnold, Brezzi, and Fortin [51], a continuous linear/linear element where the displacement interpolation is enhanced with a cubic bubble function (*i.e.*, the enhancement vanishes on the boundary). Another example is a member of the Crouzeix-Raviart element family [52], which uses continuous quadratic displacement interpolation together with a cubic bubble function and discontinuous linear pressure.

Figure 2.5: Element types and native *inf-sup* satisfaction

[50]: Donea et al. (2003), *Finite Element Methods for Flow Problems*

[51]: Arnold et al. (1984), “A stable finite element for the Stokes equations”

[52]: Crouzeix et al. (1973), “Conforming and nonconforming finite element methods for solving the stationary Stokes equations I”

Stabilization As demonstrated, there are several stable mixed finite element formulations. However, it can be cumbersome to prove any possible combination of orders of shape functions—a task that gets even bigger when considering other mixed formulations, including three-field formulations. Furthermore, the possibility of using arbitrary combinations of polynomial orders is desirable, as this opens up the possibility of choosing the order of shape functions depending on other factors posed by the specific problem at hand. Actively stabilizing the mixed finite element formulation can achieve this.

Hughes [53] proposed the sub-grid approach to stabilize mixed finite element formulations that do not satisfy the *inf-sup* condition. Later, the *variational multiscale method* (VMS) [54] was presented as a paradigm for subgrid-scale modules. The underlying assumption is that the solution of the problem in its non-discretized form consists of two parts: one that *can* be captured by the finite element approximation and one that *cannot*. The idea of the VMS approach is to enlarge Galerkin's space of approximation by a space of finer resolution (scale) and, since the missing component cannot be calculated, to approximate its effect.

Following this approach, the unknowns are expanded by the so-called sub-grid scales such that

$$\mathbf{u} \simeq \mathbf{u}_h + \tilde{\mathbf{u}} \quad (2.64)$$

$$p \simeq p_h + \tilde{p}, \quad (2.65)$$

and, in case of the three-field $u/p/e$ formulation

$$\mathbf{e} \simeq \mathbf{e}_h + \tilde{\mathbf{e}}. \quad (2.66)$$

The approximation of the subscales is obtained by introducing Eqs.(2.64–2.66) into the corresponding weak forms of the respective problems, Eqs. (2.42) and (2.57). The subscales can then be expressed in terms of the projections of the residuals of the finite element components onto the finer solution space. For the two-field u/p problem this concludes to

$$\tilde{\mathbf{u}} \approx \tau_u \tilde{\mathbf{P}}_u [\mathbf{R}_u^h] = \tau_u \tilde{\mathbf{P}}_u [\nabla \cdot (\mathbb{C}' : \nabla^s \mathbf{u}_h) + \nabla p_h + \mathbf{b}] \quad (2.67)$$

$$\tilde{p} \approx \tau_p \tilde{\mathbf{P}}_p [R_p^h] = \tau_p \tilde{\mathbf{P}}_p [\nabla \cdot \mathbf{u}_h - p_h/\kappa] \quad (2.68)$$

and for the three-field $u/p/e$ formulation the subscales are approximated as

$$\tilde{\mathbf{u}} \approx \tau_u \tilde{\mathbf{P}}_u [\mathbf{R}_u^h] = \tau_u \tilde{\mathbf{P}}_u [\nabla \cdot (\mathbb{C}' : \mathbf{e}_h) + \nabla p_h + \mathbf{b}] \quad (2.69)$$

$$\tilde{p} \approx \tau_p \tilde{\mathbf{P}}_p [R_p^h] = \tau_p \tilde{\mathbf{P}}_p [\nabla \cdot \mathbf{u}_h - p_h/\kappa] = (2.68) \quad (2.70)$$

$$\tilde{\mathbf{e}} \approx \tau_e \tilde{\mathbf{P}}_e [\mathbf{R}_e^h] = \tau_e \tilde{\mathbf{P}}_e [\mathbb{D} : \nabla^s \mathbf{u} - \mathbf{e}], \quad (2.71)$$

where τ_u , τ_p and τ_e are the *stabilization parameters* and $\tilde{\mathbf{P}}_u$, $\tilde{\mathbf{P}}_p$ and $\tilde{\mathbf{P}}_e$ are the *projection operators*, which can take different shapes.

The simplest choice is to take the projection operator as the 2nd order unity tensor $\tilde{\mathbf{P}}_u = \tilde{\mathbf{P}}_p = \tilde{\mathbf{P}}_e = \mathbf{I}$. This method is referred to as the *algebraic sub-grid scales* method or ASGS [55]. A more complex approach is based on selecting the space of the unknown sub-grid solution orthogonal to the finite element space. In this case, the projections are taken as $\tilde{\mathbf{P}}_u = \tilde{\mathbf{P}}_p = \tilde{\mathbf{P}}_e = \mathbf{P}_h^\perp = \mathbf{I} - \mathbf{P}_h$, where $\mathbf{P}_h$ denote a projection onto the finite element space. This approach is referred to as *orthogonal subscale stabilization* or OSS [56].

The choice of stabilization parameter [57] impacts the obtained convergence rates of the different unknowns. The lengths-scales used and whether they depend on the mesh size (and therefore decrease upon mesh refinement) are of particular interest. Tab. 2.3 summarizes the different stabilization parameter choices, where $C^\circ = (1/2G + 2/3\kappa)^{-1}$

[53]: Hughes (1995), “Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods”

[54]: Hughes et al. (1998), “The variational multiscale method—a paradigm for computational mechanics”

[55]: Codina (2000), “On stabilized finite element methods for linear systems of convection–diffusion–reaction equations”

[56]: Codina (2000), “Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods”

[57]: Badia et al. (2009), “Unified Stabilized Finite Element Formulations for the Stokes and the Darcy Problems”

is the compressibility modulus that reduces to $C^\circ = 2G$ in the incompressible limit ($\kappa \rightarrow \infty$). The stabilization parameters originate from a Fourier analysis of the problem for the subscales and depend on h (element size), L_0 (characteristic length of geometry) or a combination of both [35, 57].

Tab. 2.4 summarizes the obtained convergence rates for the different primary unknowns for a linear/linear/linear $u/p/e$ formulation.

Parameter	h^2 -Stab.	hL_0 -Stab.
τ_u	$c_u h^2 / 2G$	$c_u h L_0 / 2G$
τ_p	$c_p \tilde{C}^\circ$	$c_p \tilde{C}^\circ h / L_0$
τ_e	c_e	$c_e h / L_0$

Error	h^2 -Stab.	hL_0 -Stab.
Global $\mathbf{u}$	2.0	1.5
Global p	1.0	1.5
Global $\mathbf{e}$	1.0	1.5
Local $\mathbf{u}$	1.0	0.5
Local p	—	0.5
Local $\mathbf{e}$	—	0.5

[35]: Chiumenti et al. (2021), “Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation”

[57]: Badia et al. (2009), “Unified Stabilized Finite Element Formulations for the Stokes and the Darcy Problems”

Table 2.3: Different stabilization approaches

Table 2.4: Global and local convergence rates for displacement, pressure and strain for different stabilization approaches (linear/linear/linear elements)

Projections The idea of VMS stabilization is to enlarge Galerkin’s space of approximation by a space of finer resolution, which allows the approximation of the part of the continuous solution that the finite element approximation cannot capture. These subscales are based on the residuals of the finite element components, which are computed at the element’s integration points.

If the subscales are directly used to build the stabilized element stiffness matrix without further modification, the method is referred to as ASGS. However, it has proven to be beneficial to only use those parts of the subscales that belong to a space orthogonal to that of the finite element approximation. This approach was introduced as OSS. To extract the orthogonal component of the subscales, they are projected onto the finite element space and this part is then subtracted from the original subscales approximation. The remaining subscales are therefore entirely defined in a space orthogonal to that of the finite element approximations.

From an implementational point of view, this operation requires projecting the subscales from the integration points (where they are computed) to nodes (where the finite element approximation is defined), which effectively projects them onto the space of the finite element approximation. The projection can either be carried out by assembling and solving a linear system, including a consistent mass matrix, or by performing a lumped mass matrix projection, which is less computationally expensive, but also less accurate.

Either way, since the subscales depend on the unknowns, and said unknowns are arguments of the projection operations (see Eq. (2.67–2.71)), it becomes difficult to isolate the unknowns in order to solve the linear system. In practice, one of four approaches can be taken:

1. In the case of ASGS, the subscales are used directly without projecting them first. In this case, the unknowns can be isolated easily, and the linear problem can be solved straightforwardly. If the mechanical problem was linear initially, the system remains linear, and the solution can be obtained without iterating.

In the case of OSS, the projections onto the finite element space prevent an isolation of the unknowns.

2. The projections of the residuals can be treated as unknowns themselves. However, it severely increases the number of unknowns, which becomes unsustainable quickly.
3. The projections can be computed using the previous iteration's solution. This approach is also very accurate but does not increase the number of unknowns further. It, however, forces the solver to iterate several times to reach equilibrium.
4. The projections can be computed using the previous step's solution. This is the least accurate approach. However, assuming that the subscales only marginally change between time steps, this is a good approach since no (additional) non-linearity is introduced to the system.

Stabilized two-field u/p formulation The fully stabilized u/p formulation is obtained after introducing the scale-split approximations of the unknowns, stated in Eq. (2.68) into the weak form of the mixed u/p problem, stated in Eq. (2.42). The Galerkin discrete version of the stabilized problem reads

$$\begin{aligned}
& \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV + \int_{\Omega} \nabla \cdot \delta \mathbf{u}_h \, p_h \, dV \\
& + \int_{\Omega} \delta p_h \, \nabla \cdot \mathbf{u}_h \, dV - \int_{\Omega} \delta p_h \, p_h / \kappa \, dV \\
& + \sum_K \int_{K(\Omega)} -\nabla \delta p_h \cdot \tilde{\mathbf{u}} \, dV \\
& + \sum_K \int_{K(\Omega)} (-\nabla \cdot \delta \mathbf{u}_h + \delta p_h / \kappa) \, \tilde{p} \, dV \\
& = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA.
\end{aligned} \tag{2.72}$$

Remark 2.4.1. Assuming linear shape function, a selection of stabilization terms can be ignored such as $\sum_K \int_{K(\Omega)} -\nabla^s \delta \mathbf{u} : \mathbb{C}' : \nabla^s \tilde{\mathbf{u}} \, dV$ since they would include 2^{nd} order derivatives of the linear approximation of unknowns.

A key property of the OSS stabilization is the orthogonal projections onto the finite element space, which keep the consistency of the formulation in a weak sense despite only including a minimum number of terms to stabilize the solution (see [58]). For this particular formulation only the term $\sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV$ is required to achieve stability.

Following this approach, the final form is obtained. Further mentions of the *stabilized u/p formulation* refer to this equation.

[58]: Codina et al. (2017), “Variational Multiscale Methods in Computational Fluid Dynamics”

$$\begin{aligned}
& \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV + \int_{\Omega} \nabla \cdot \delta \mathbf{u}_h p_h \, dV \\
& + \int_{\Omega} \delta p_h \nabla \cdot \mathbf{u}_h \, dV - \int_{\Omega} \delta p_h p_h / \kappa \, dV \\
& - \sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA.
\end{aligned} \tag{2.73}$$

$$\begin{aligned}
& B^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) \\
& B_{\text{oss}}^{\text{up}}(\delta p_h, p_h) = L(\delta \mathbf{u}_h)
\end{aligned} \tag{2.74}$$

Stabilized three-field $u/p/e$ formulation The fully stabilized $u/p/e$ formulation is obtained after introducing the scale-split approximations of the unknowns, stated in Eq. (2.71), into the weak form of the mixed $u/p/e$ problem, stated in Eq. (2.57). The Galerkin discrete version of the stabilized problem reads

$$\begin{aligned}
& \int_{\Omega} \nabla \cdot \delta \mathbf{u}_h p_h \, dV + \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \mathbf{e}_h \, dV \\
& + \int_{\Omega} \delta p_h \nabla \cdot \mathbf{u}_h \, dV - \int_{\Omega} \delta p_h p_h / \kappa \, dV \\
& + \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV - \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C} : \mathbf{e}_h \, dV \\
& + \sum_K \int_{K(\Omega)} (\mathbb{C}' : \nabla \delta \mathbf{e}_h - \nabla \delta p_h) \cdot \tilde{\mathbf{u}} \, dV \\
& + \sum_K \int_{K(\Omega)} (-\nabla \cdot \delta \mathbf{u}_h + \delta p_h / \kappa) \tilde{p} \, dV \\
& + \sum_K \int_{K(\Omega)} (\nabla^s \delta \mathbf{u}_h : \mathbb{C}' + \delta \mathbf{e}_h : \mathbb{C}) : \tilde{\mathbf{e}} \, dV \\
& = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA.
\end{aligned} \tag{2.75}$$

As for the u/p formulation, the necessary stabilization terms can be reduced, either because they vanish, assuming linear elements or because they are not required to achieve stability (requires OSS). For this particular formulation, only the terms $\sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV$ and $\sum_K \tau_e \int_{K(\Omega)} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \tilde{\mathbf{P}}_h^\perp [\mathbb{D} : \nabla^s \mathbf{u}_h] \, dV$ are required to achieve stability.

Following this approach, the final form is obtained. Further mentions of the *stabilized $u/p/e$ formulation* refer to this equation.

$$\begin{aligned}
& \int_{\Omega} \nabla \cdot \delta \mathbf{u}_h p_h \, dV + \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \mathbf{e}_h \, dV \\
& + \int_{\Omega} \delta p_h \nabla \cdot \mathbf{u}_h \, dV - \int_{\Omega} \delta p_h p_h / \kappa \, dV \\
& + \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV - \int_{\Omega} \delta \mathbf{e}_h : \mathbb{C} : \mathbf{e}_h \, dV \\
& + \sum_K \tau_e \int_{K(\Omega)} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \mathbf{P}_h^\perp [\mathbb{D} : \nabla^s \mathbf{u}_h] \, dV \\
& - \sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA.
\end{aligned} \tag{2.76}$$

$$\begin{aligned}
& B^{\text{upe}}([\delta \mathbf{u}_h, \delta p_h, \delta \mathbf{e}_h], [\mathbf{u}_h, p_h, e_h]) \\
& + B_{\text{oss}}^{\text{upe}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) = L(\delta \mathbf{u}_h)
\end{aligned} \tag{2.77}$$

2.5 Adaptive Formulation Refinement

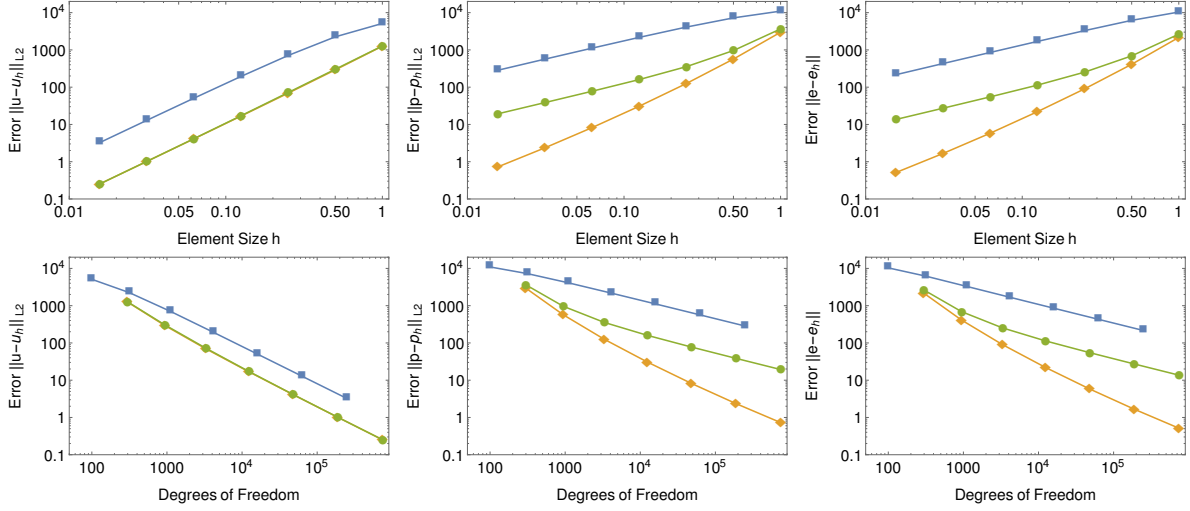
The three-field $u/p/e$ formulation is a powerful tool to approach the incompressible limit, address locking in nearly incompressible cases and guarantee local convergence of stresses and strains. However, there are some drawbacks which explain why the formulation is not used as the standard in finite element simulations.

2.5.1 Motivation

Ultimately, the additional accuracy originates from the increased number of unknowns. Tab. 2.5 shows the number of unknowns per node and the number of entries in element stiffness matrices of different element types. Higher numbers of nodal DOF increase the size of the linear system and lead to increased solution times.

	Formulation	DOF per Node	Entries $\mathbf{K}_{\text{tri}}$	Entries $\mathbf{K}_{\text{quad}}$
2D	standard FE	2	36	64
	u/p	3	81	144
	$u/p/e$	6	324	576
	Formulation	DOF per Node	Entries $\mathbf{K}_{\text{tet}}$	Entries $\mathbf{K}_{\text{hex}}$
3D	standard FE	3	144	576
	u/p	4	256	1024
	$u/p/e$	10	1600	6400

Table 2.5: Number of degrees of freedom and size of elemental stiffness matrices for tetrahedral and hexahedral elements.



The trade-off between computational effort and accuracy for the u/p and $u/p/e$ formulations can be assessed through a mesh convergence study. The error norms are computed as follows.

$$\|u - u_h\|_{L^2} = \left(\int_{\Omega} (u - u_h) \cdot (u - u_h) dV \right)^{1/2} \quad (2.78)$$

$$\|p - p_h\|_{L^2} = \left(\int_{\Omega} (p - p_h) \cdot (p - p_h) dV \right)^{1/2} \quad (2.79)$$

$$\|e - e_h\| = \left(\int_{\Omega} (s - s_h) : (e - e_h) dV \right)^{1/2} \quad (2.80)$$

The error of the displacement and pressure are evaluated in the global L_2 norm, whereas the third error is the energy error, which is evaluated in the global energy error norm, which, assuming linear elasticity, gives identical convergence rates as the deviatoric strain error in the L_2 norm.

Plotting the errors relative to the element size h and the number of degrees of freedom reveals clear trends. For *any* element size (see Fig. 2.6, top), the error of *every* unknown is lower for the $u/p/e$ formulation than for the u/p formulation. This is to be expected since adding the additional unknowns increases the accuracy of the formulation. However the analysis must be done with respect to the number of DOFs, since this is a much better indication of the computational effort than the element size h . Fig. 2.6 (bottom) clearly shows that for a given number of DOF, the error is always lowest for the $u/p/e$ formulation – or, for a given target error, fewer DOFs are required, *i.e.*, a coarser mesh is sufficient. This effect is even stronger when the convergence rate is increased (see $\blacklozenge$ $u/p/e$ hL_0 -stabilization).

The study of convergence rates shows the superiority of the three-field $u/p/e$ formulation. However, similarly to a very fine mesh (h -refinement) or higher order interpolation functions (p -refinement), there might be areas of the computational domain that do not require a very accurate solution. The most common case might be when the solution is very smooth in one part of the computational domain. Adaptive refinement schemes are widely used to balance accuracy and computa-

Figure 2.6: Global error of displacements, pressure and deviatoric strain over element length h (top) and degrees of freedom (bottom), markers are $\blacksquare$ u/p , $\bullet$ $u/p/e$ (h^2 -stabilization) and $\blacklozenge$ $u/p/e$ (hL_0 -stabilization)

tional efficiency.

2.5.2 Methodology

The proposed *adaptive formulation refinement* method aims to provide flexibility when using mixed formulations to dynamically switch between the u/p and $u/p/e$. The used refinement criteria can be geometrically predefined, based on physical parameters, or based on a posteriori error estimation.

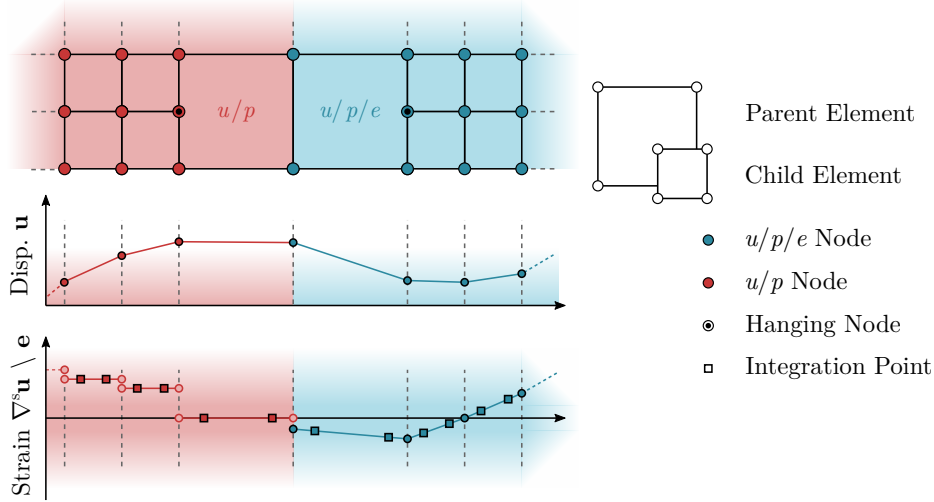


Figure 2.7: Displacement and strain interpolation at AFR interface.

At the start of each time step, the refinement criterion is evaluated, defining the u/p and $u/p/e$ subdomains. Element stiffness matrices are then computed according to the subdomain to which each element belongs.

At the interface between u/p and $u/p/e$ elements, all three-field variables (u , p , and e) are retained. The interface between two-field and three-field elements is shown in Fig. 2.7, where all nodes at the interface are $u/p/e$ nodes.

In the u/p subdomain, the strain is calculated by computing the symmetric gradient of the displacement (which is C^0 -continuous) at the integration points, resulting in a C^{-1} -continuous (*i.e.*, discontinuous) strain field. In contrast, within the $u/p/e$ subdomain, the strain is directly obtained as a nodal solution and then interpolated to the integration points, resulting in a C^0 -continuous strain field. This setup allows for the stresses and energy dissipation to be computed at the integration points, regardless of the subdomain.

Remark 2.5.1. *From an implementational point of view, it is beneficial to allocate the entire linear system and the element stiffness matrices as $u/p/e$ and deactivate the corresponding, strain-related terms if the element has been flagged as u/p .*

Remark 2.5.2. *Special attention must be paid to the computation of the strain subscales $\tilde{\mathbf{e}}$ when using orthogonal sub-grid stabilization (OSS). The computation of the projection onto the space of subscales $\tilde{\mathbf{P}}$*

involves the nodal projection of the subscales of the affected variables. This poses no problem for the displacement and pressure subscales since both fields are defined on the whole domain. However, the strain subscales only exist in the $u/p/e$ subdomain. Consequently, the nodes at the interface lack a contribution from adjacent u/p elements when the nodal projection of the strain subscales is performed. As a remedy, the nodal strain in the u/p subdomain is approximated as a nodal projection of the symmetric displacement gradient.

2.5.3 Combination with h -refinement

Adaptive formulation refinement adds additional degrees of freedom, depending on the accuracy required in specific regions of the computational domain. However, only a single “level of refinement” can be applied. If the $u/p/e$ formulation, in conjunction with the current mesh size, is not sufficiently accurate, further mesh refinement is required.

The combination of adaptive formulation and mesh refinement is straightforward and requires no special adjustments. The process begins with h -refinement, during which the mesh is refined, the solution fields are interpolated onto the new mesh and hanging nodes are addressed. Subsequently, the AFR refinement criterion is *re-evaluated* on the refined mesh to determine the optimal u/p and $u/p/e$ subdomains. Recalculating the domain partitioning yields better results than relying on the *interpolation* of the previous partitioning information onto the refined mesh.

2.5.4 Refinement Criteria

However, the development of suitable refinement criteria is a more intricate task. Refinement criteria can be divided into two main categories: a priori and a posteriori. The former is not very practical in most cases, as it requires prior knowledge of where the mesh and/or formulation need refinement – information that is typically unavailable in real-world applications. By contrast, a posteriori criteria are far more practical. These criteria rely on data derived from the simulation results to determine where refinement is required. A posteriori criteria can also be classified into several distinct groups.

Error estimation based Error Estimators (EE) serve as a valuable tool to locate regions in need of refinement. The challenge lies in creating a metric that reliably approximates the local error without knowledge of the exact solution. A critical requirement for such estimators is that the predicted error must diminish according to the refinement in a manner consistent with the reduction of the actual error.

One example is the well-known Zienkiewicz-Zhu (ZZ) error estimator ([59]), which compares the discontinuous strain field $\mathbf{e}_d = \nabla^s \mathbf{u}|_{\xi=0}$ with its projection onto the space of finite element functions (thus continuous) $\mathbf{e}_c = P_h [\nabla^s \mathbf{u}]|_{\xi=0}$, as the latter one is assumed to be a more

[59]: Zienkiewicz et al. (1987), “A simple error estimator and adaptive procedure for practical engineering analysis”

accurate measure of the strain. Typically the comparison is carried out at the element's center ($\boldsymbol{\xi} = \mathbf{0}$).

$$E = \int_{\Omega^e} (\mathbf{e}_c - \mathbf{e}_d) : (\mathbf{e}_c - \mathbf{e}_d) dV \quad (2.81)$$

For linear, C^0 -displacement-continuous elements, the ZZEE approximately quantifies the difference between the strain field obtained by computing the symmetric gradient of the displacement field in the u/p formulation and the strain field that would be obtained using the $u/p/e$ formulation (which is naturally C^0 -continuous). It is therefore a good choice to drive the formulation change.

A different approach is based on the VMS framework, used to stabilize mixed finite element formulation. The subscales were introduced as the part of the continuous solution that the finite element approximation could not capture and are approximated using the projection of the residuals of the finite element components onto the finer solution space, see Eqs. (2.67-2.71). It turns out that these estimated subscales can be utilized as error estimators. Baiges and Codina [60] demonstrated that the proposed error estimator accurately approximates the energy norm error and is effective for both simple linear constitutive models and complex nonlinear cases. VMS-based error estimators have been used in multiple applications such as anisotropic fluid mechanics simulations [61] or compressible flow simulations [62].

The estimated error is ultimately compared against user-defined error threshold values. It is good practice to define two criteria to avoid oscillating around the threshold. A robust refinement criterion should have an upper refinement threshold $E_{\max}$ and a lower de-refinement threshold $E_{\min}$, with $E_{\max} > E_{\min}$.

$$\begin{aligned} E &\geq E_{\max} && \rightarrow \text{refinement} \\ E &< E_{\min} && \rightarrow \text{de-refinement} \end{aligned} \quad (2.82)$$

Physics based Physics-based refinement criteria utilize problem-specific measures to identify regions requiring refinement. For instance, in J_2 plasticity, plastic (i.e., isochoric) strains develop in certain parts of the computational domain. In these areas, a FE formulation capable of handling isochoric deformations must be employed. Similarly, in fracture mechanics, the crack tip is of particular interest, necessitating an enhanced strain/stress accuracy formulation near the crack tip to mitigate mesh bias. In both scenarios, the advanced formulation is activated when approaching the yield surface. In the context of plasticity this means switching from $u \rightarrow u/p$ and in fracture mechanics it means switching from $u \rightarrow u/\varepsilon$.

A useful refinement criterion can be derived from the ratio R between the equivalent stress σ_{eq} and the material property that dictates the onset of inelastic deformation. For plasticity, this is the yield stress ($R = \sigma_{\text{eq}}/\sigma_Y$), while for damage models (assuming mode I cracking), the tensile strength ($R = \sigma_{\text{eq}}/f_T$) is a good choice. Threshold values

[60]: Baiges et al. (2017), “Variational Multiscale error estimators for solid mechanics adaptive simulations: An Orthogonal Subgrid Scale approach”

[61]: Bazile et al. (2018), “Variational Multiscale error estimator for anisotropic adaptive fluid mechanic simulations: Application to convection–diffusion problems”

[62]: Bayona-Roa et al. (2018), “Variational multiscale error estimators for the adaptive mesh refinement of compressible flow simulations”

($0 < \alpha_{\text{de-ref}} < \alpha_{\text{ref}} < 1$) can then be used to define the refinement criterion, determining where refinement or de-refinement is necessary.

$$\begin{aligned} R &\geq \alpha_{\text{ref}} && \rightarrow \text{refinement} \\ R &< \alpha_{\text{de-ref}} && \rightarrow \text{de-refinement.} \end{aligned} \tag{2.83}$$

Remark 2.5.3. *A third category are geometry based refinement criteria but they play a subordinate roll regarding formulation refinement and are discussed instead in Sec. 3.3 in the context of embedded domain methods.*

Multi-target criteria The challenge in combining multiple refinement methods (*e.g.*, formulation and mesh refinement) lies in designing a unified criterion that governs the refinement and de-refinement of both methods. The selection of refinement and de-refinement thresholds is critical, as it determines the prioritization between the two approaches. Given that adaptive formulation refinement (AFR) provides only a single “level of refinement”, it is logical to set its thresholds lower than those of adaptive mesh refinement (AMR). This approach ensures that the enhanced accuracy of the mixed formulation is utilized first. If the desired accuracy is still not achieved, mesh refinement is then applied as an additional measure.

2.5.5 Solver Strategies

The driving idea of the adaptive formulation refinement approach is to be able to deactivate the degrees of freedom (DOFs) associated with the strain in the u/p subdomain to reduce memory requirements and computational time. However, special attention must be given to the structure of the system matrices.

In large numerical simulations, matrices are typically stored in a *Compressed Sparse Row* (CSR) format. The reason for the sparseness of the matrix is that FE nodes only interact with their immediate neighbors. When ordering the nodes in a beneficial manner, only a narrow band of non-zero entries appears in the matrix, while the overwhelming majority of the matrix is empty. To avoid storing the empty parts of the matrix, the CSR format relies on a matrix graph to track the non-zero entries only.

To further enhance the efficiency of the CSR format, the graph can be “built by nodes” rather than “by DOF”. Typically, within a given element, all DOFs of a particular node are connected to all other DOFs of all other nodes. Therefore, it makes sense to consider all DOFs of a node as a whole. Hence, in many Finite Element applications, it is common to build the graph by nodes. This means that all DOFs associated with a node are treated as non-zeroes and assembled together. This approach suits standard cases where all nodes have the same number of DOFs.

However, for the AFR method, the strain components in the u/p subdomain are not utilized, resulting in zero entries in the corresponding positions of the elemental stiffness matrices. Without adjustments, the

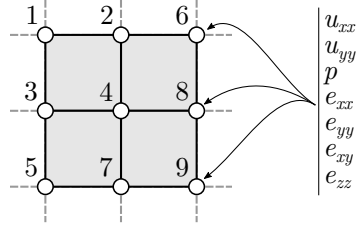


Figure 2.8: Four elements with potentially 7 DOF per node

graph would mirror that of the default $u/p/e$ formulation, leading to the unnecessary assembly of numerous zeroes into the system matrix. This not only slows down matrix build times but also impacts solve times, as the solver initially treats these entries as non-zero due to their presence in the matrix graph.

To address this, the graph must be constructed by DOFs rather than nodes. This approach includes the displacement and pressure DOFs in the u/p subdomain, along with a unitary value on the main diagonal for the strain components. Crucially, all off-diagonal entries in the rows and columns associated with the strain DOFs are completely removed from the graph.

As a result, the unitary values on the main diagonal become the only non-zero entries for these rows and columns, which significantly improves the solving times. The optimization also preserves a consistent matrix size, eliminating the need for renumbering DOFs. The activation and deactivation of DOFs are managed at runtime.

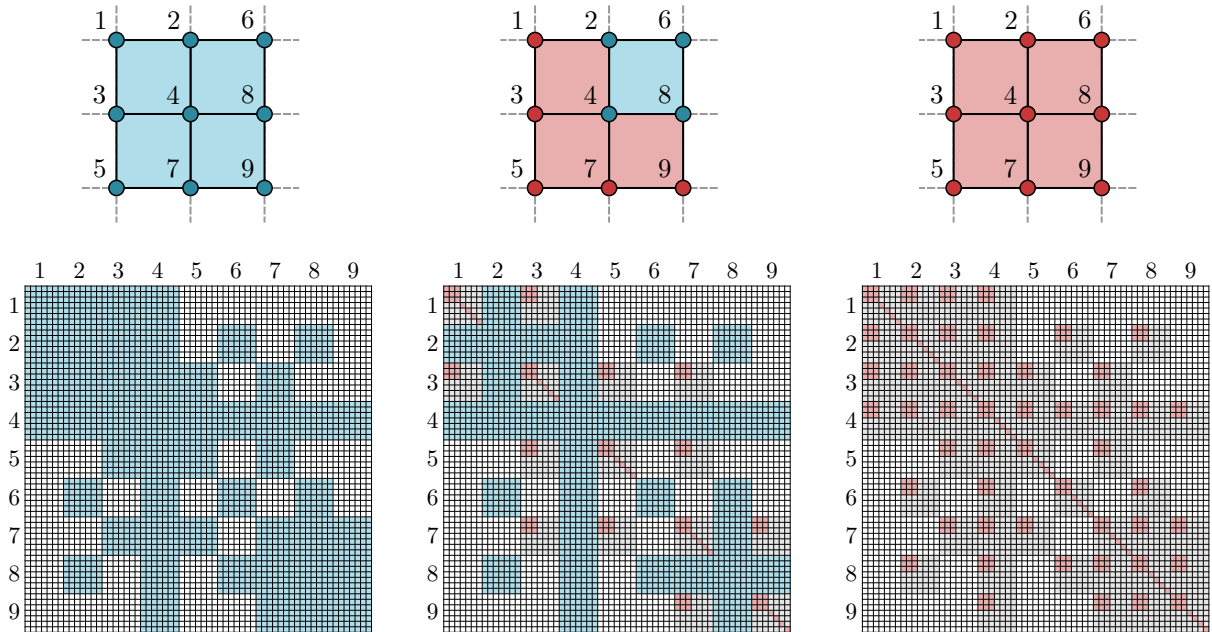


Fig. 2.9 illustrates the zero and non-zero values of the system matrix for an example case of four elements with seven DOF per node (Fig. 2.8). Zero entries are represented by white and gray squares and non-zero entries are represented by red and blue squares. When building the system by nodes, the gray entries (although being zeroes) are assembled to the system regardless, slowing down build and solve times. When building the system by DOF, they are eradicated from the graph, thus speeding up build and solve times.

Figure 2.9: Four elements with potentially 7 DOF per node

Remark 2.5.4. *In practice, the PETSC library ([63]) is used to construct and solve the linear system of equations arising from the FE analysis. In PETSC, a ‘BAIJ-matrix’ (B for ‘blocked’) allows for multiple DOFs to be assembled as one block in a single operation (i.e., graph built by nodes). To allow a variable number of DOFs per node, the ‘AIJ-matrix’ option must be chosen in PETSC.*

[63]: Balay et al. (2021), *PETSc/TAO Users Manual*

2.5.6 Results

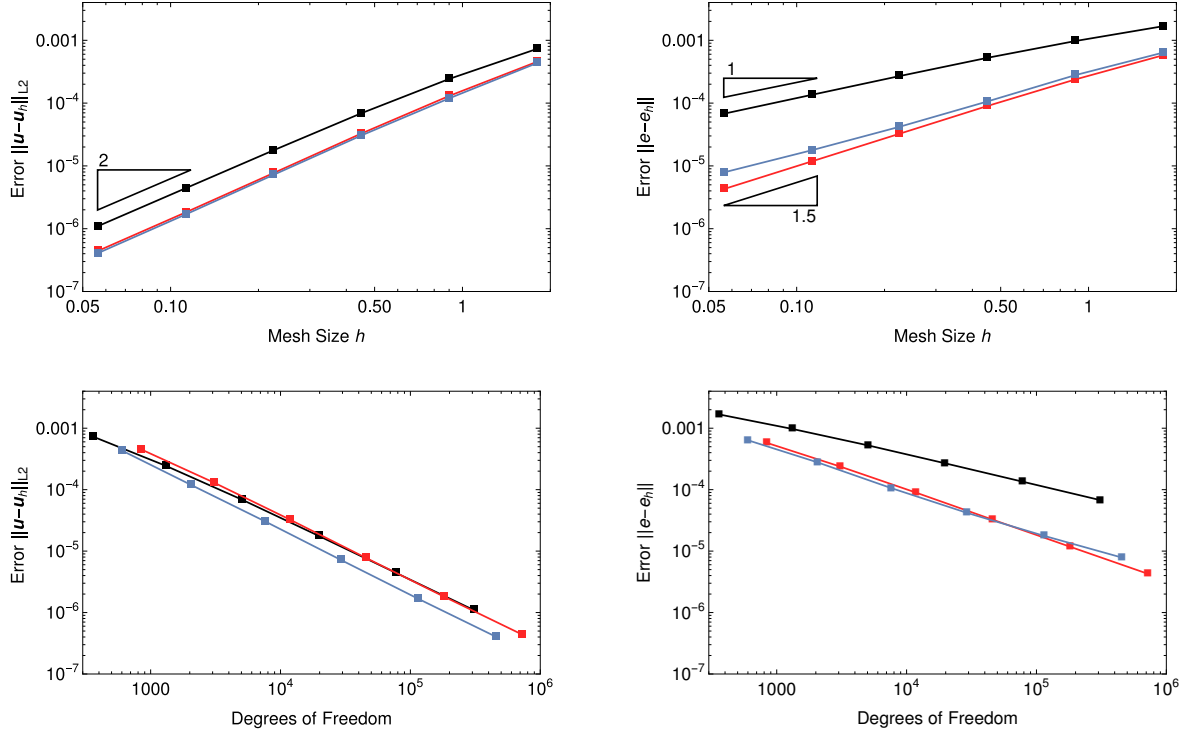
Convergence Study Sec. 2.5 showed the convergence rate of the $u/p/e$ formulation, compared to the u/p one. This has motivated the development of the AFR approach. Fig. 2.10 shows the optimal convergence rate obtained using the AFR method with respect to the used DOF. The detailed description of the problem is omitted at this point. See Venghaus et al. [64][Appx. A]

It can be clearly seen that the displacement error (Fig. 2.10, left) decreases with the same convergence rate regardless of the formulations. Thus, for a given mesh size h , the $u/p/e$ (■) and AFR (■) solutions yield similar results (Fig. 2.10, top-left). However, when the error is displayed over the number of DOFs (Fig. 2.10, bottom-left), the AFR (■) solution always yields the highest accuracy for a given number of DOFs.

[64]: Venghaus et al. (2023), “An accurate approach to simulate friction stir welding processes using adaptive formulation refinement”

This effect can only be achieved when the computational domain is reasonably partitioned into u/p and $u/p/e$ subdomains. The principle guiding adaptive formulation refinement mirrors that of mesh refinement: the refinement must be directed toward regions that derive the greatest benefit, which, in this case, means areas where the three-field formulation is most advantageous. When implemented effectively, the AFR approach can achieve the same level of accuracy as a uniformly applied three-field formulation but at a reduced computational cost.

A slightly different behavior can be observed when the energy error (i.e., deviatoric strain error) is studied (Fig. 2.10, right). The $u/p/e$ (■) formulation yields both the highest convergence rate and the most accurate results. In contrast, the u/p (■) formulation yields a lower convergence rate and less accurate results. The AFR solution initially follows the solution of the $u/p/e$ formulation (large h) and, therefore, performs marginally better when displayed over DOF. However, for smaller h , the convergence rate aligns more with that of the u/p (■) formulation, although still yielding much more accurate results.



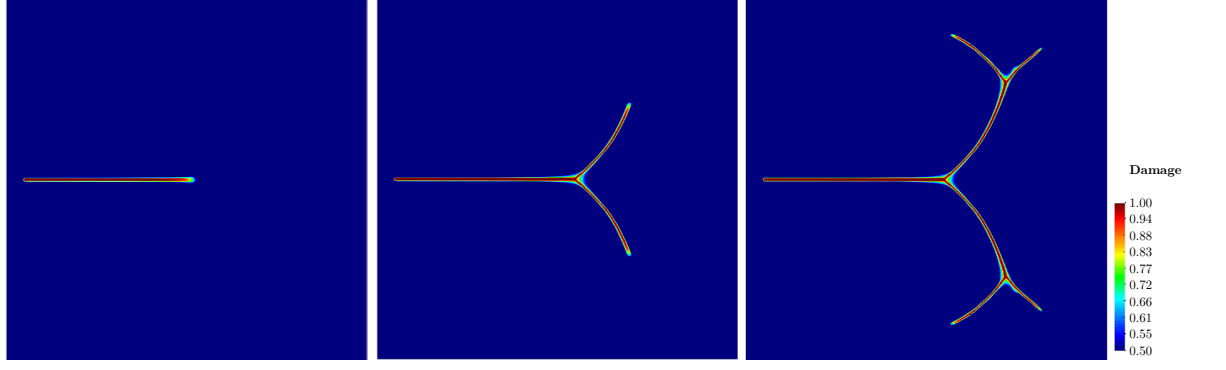
A possible explanation for this phenomenon is that, as the mesh gets finer, and the solution in the $u/p/e$ subdomain converges with a rate of 1.5 (as expected, see Tab. 2.4), the solution in the u/p subdomain only converges with a rate of 1.0. As h further decreases, the error in the u/p subdomain begins to dominate, causing the overall convergence rate to gradually decrease from 1.5 to 1.0. This effect could be countered or prolonged by increasing the size of the $u/p/e$ subdomain. Ultimately, the same considerations must be made as if adaptive mesh refinement was used. AMR aims to enhance the versatility of finite element formulations and tackle problems that cannot be addressed by uniformly refining the mesh. AMR is not used to prove convergence rates. Similarly, the intent of the AFR approach is to allow the use of mixed FE formulations in cases where it would be unfeasible to use the more computationally expensive formulation everywhere.

Cracking Application The following example demonstrates how AFR and AMR can be used together to tackle cracking phenomena. The detailed description of the problem is omitted at this point. See Barbat et al. [65][Appx. B] for more details.

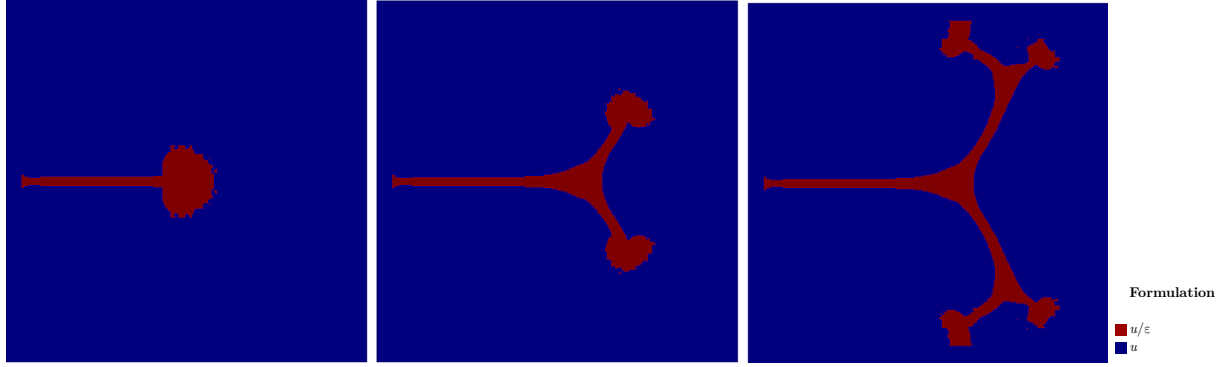
The results depicted in Fig. 2.12 show the interaction of mesh and formulation refinement. Physics-based refinement criteria were used for both refinement strategies. The thresholds for refinement and de-refinement are listed in Tab. 2.6.

Figure 2.10: Global error of displacements (left) and deviatoric strain (right) over element length h (top) and degrees of freedom (bottom), markers are: ■ u/p formulation, ■ $u/p/e$ formulation and ■ AFR.

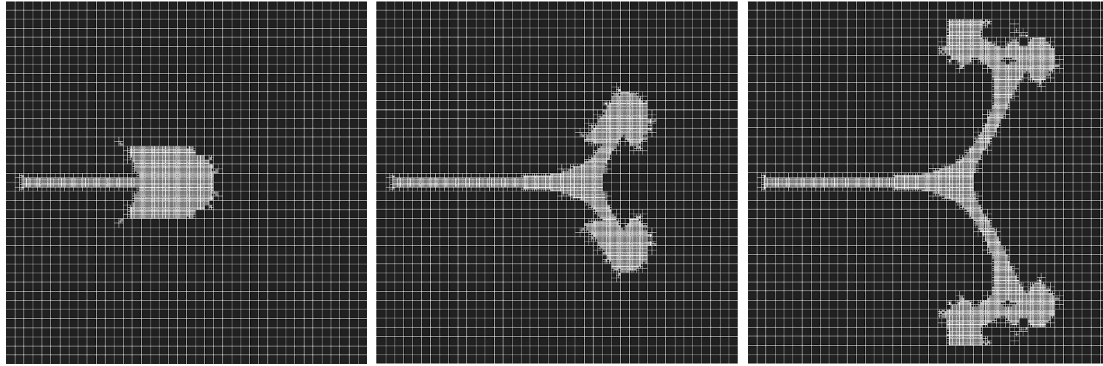
[65]: Barbat et al. (2022), “Objectivity in quasi-brittle structural failure via adaptive formulation and mesh refinement”



(a) Damage evolution



(b) Adaptive Formulation Refinement



(c) Adaptive Mesh Refinement

	AFR	AMR
refinement ($R \geq$)	$\alpha_{\text{ref}} = 0.75$	$\beta_{\text{ref}} = 0.75$
de-refinement ($R <$)	$\alpha_{\text{de-ref}} = 0.50$	$\beta_{\text{de-ref}} = 0.50$

Figure 2.12: Crack evolution in branching test at different stages**Table 2.6:** Refinement and de-refinement thresholds

The refinement threshold is typically higher than the de-refinement threshold to avoid oscillations. Here both AFR and AMR use identical values. Consequentially the refined subdomains look very similar. Both mesh and formulation are refined at the crack tip, where the highest stresses occur. After the crack progressed and the stresses are relaxed, the mesh and formulation are de-refined. However, an additional constraint prevents the elements affected by the crack to return to their original, un-refined state. Thus, de-refinement is prohibited in both strategies when elements have experienced damage ($d > 0$).

It is noticeable that the fine-mesh subdomain is slightly larger than the $u/p/e$ subdomain, even though both refinement methods use identical thresholds. This phenomenon arises because the process begins with mesh refinement, as was previously discussed in Sec. 2.5.3. Once the displacements, strains, and stresses are interpolated to the newly generated nodes and elements, the formulation refinement criterion is reevaluated. The increased mesh resolution allows the AFR criterion to be evaluated more precisely, resulting in a slightly smaller $u/p/e$ subdomain since de-refinement happens at a lower threshold mesh coarsening does not catch up immediately.

2.6 Concluding Remarks

This first part of the thesis addresses the numerical difficulties that arise when dealing with isochoric material behavior. These difficulties exist in both solid and fluid mechanics, either because the material is intrinsically incompressible or because the plastic flow is isochoric. Furthermore, the necessity of strain- and stress-accurate formulations is demonstrated in the context of fracture mechanics.

Early attempts to deal with isochoric material behavior are presented, alongside an examination of the possibilities and limitations of mesh size (h) and polynomial order (p) refinement. Mixed formulations are introduced as an effective way to address isochoric behavior (two-field u/p formulation) and to boost strain and stress accuracy, ensuring local strain and stress convergence in linear elements (three-field $u/p/e$ formulation). Stability issues are addressed using the variational multiscale framework and, in particular, the method of orthogonal subscales.

Next, the idea of adaptive formulation refinement is presented as a means of applying the accurate but computationally expensive three-field formulation only in selected subdomains, balancing accuracy and efficiency. Different refinement criteria are considered, including multi-target criteria that drive both mesh and formulation refinement simultaneously. The convergence rate of the AFR method is compared to those of the u/p and $u/p/e$ formulations.

Finally the AFR method, in combination with adaptive mesh refinement, is used in the context of fracture mechanics.

Unfitted FEM / Embedded Domain Methods

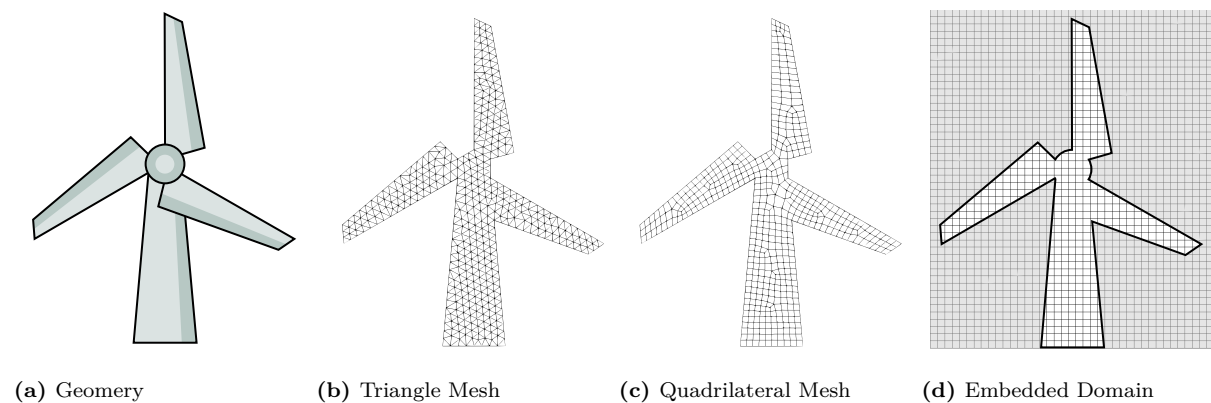
3

The second part of this work addresses complex geometries and moving boundary conditions that arise in industrial applications. Different methods are introduced, allowing for *unfitted meshes* and (material) discontinuities. A method suitable for simulating metal forming and manufacturing processes is chosen, complementing the techniques for isochoric behavior and strain and stress accuracy presented in the previous chapter.

3.1 Objectives

The *Introduction* to this work (Sec. 1) addressed the importance of numerical simulations as parts of development and production chains. Consequently, the capabilities and directions of development of these numerical methods must also be assessed in terms of their applicability to day-to-day engineering problems. Novel numerical methods are often tested on well-studied, academic benchmarks. However, the transfer of such developed methods to large-scale applications is often challenging.

The distinction between academic examples and real-world applications extends beyond the sheer size of the finite element problem, defined by the number of degrees of freedom (DOF). Efficient programming, specialized solution strategies, and optimized solvers are typically used to manage problem size. However, an equally critical yet often overlooked challenge is handling complex and highly detailed geometries. Additionally, the incorporation of intricate (sometimes dynamically changing) boundary conditions presents further difficulties that must be addressed.



The standard finite element method typically adopts body-fitted meshes, as shown in Figs. 3.1b and 3.1c. This means that the boundary of the computational domain is represented by element boundaries. However, the more complex the domain is, the less likely structured meshes can be used. The generation of unstructured meshes using tetrahedral meshes is quite robust and has been available for a long time. These

Figure 3.2: Mesh generation from geometry: body fitted and non-body fitted

meshing algorithms are typically based on Delaunay triangulation. See Green and Sibson [66] and Anglada [67] for original works and Si [68] for a modern C++ implementation. However, good quality hexahedral meshes are a much more recent development, see Shepherd and Johnson [69] and Gregson, Sheffer, and Zhang [70]. Regardless of the mesh type, the process of importing geometry (typically from CAD formats such as STEP, IGES, or STL) into a FEM software requires the use of a mesh generator. However cleaning and meshing, are laborious tasks and verifying mesh integrity and applying boundary conditions are tasks that are rarely fully automated and often require the FEM user's supervision. Matters are worse when the computational domain moves.

Another problem occurs when often performing the same simulation (meaning the same applied methods, material parameters and solution strategy) but for different geometries. As an example: designing a new product often involves an optimization loop where the CAD model must be adapted, cleaned again, and a new mesh must be generated. These steps can take up a significant portion of the entire process, even if all other simulation parameters remain unchanged.

The alternative are methods that allow for unfitted meshes (see Sec. 3.2.6).

All these methods share the same idea: the boundary of a physical domain, or a discontinuity within a physical domain (such as a crack), no longer needs to align with the mesh. Although there exist a large number of different methods, they can be grouped into two categories:

- Methods that typically use body-fitted meshes to discretize the physical domain but allow for extremely high gradients or discontinuities to appear within the elements. In solid mechanics, these could be cracks, shear bands, inclusions, or voids; in fluid mechanics, shocks, and boundary layer effects. Special enrichment functions are then used to capture these gradients or discontinuities.

These methods are generally known as *extended* or *generalized finite element methods* (XFEM/ GFEM).

- Methods where the simulation is performed on a background mesh that is typically structured. The physical object is then embedded into this domain. The embedded boundary might separate materials with different properties (*e.g.*, a stiff inclusion in a soft matrix) or describes voids in the material. However, in some cases the concept of embedding is primarily employed to facilitate handling complex geometries, see Fig. 3.1d.

Examples of this group are the *fictitious domain method* (FDM), the *immersed boundary method* (IBM), the *finite cell method* (FCM) or *Cut-FEM*.

On the one hand, methods using unfitted meshes alleviate some of the restrictions of the standard finite element method. On the other hand, they create a set of new challenges. The different methods mentioned above originated from distinct approaches to solving these new challenges. However, nowadays, these methods are partially overlapping. Therefore, the following section focuses on discussing the challenges and possible solutions rather than specific methods. After establishing

[66]: Green et al. (1978), “Computing Dirichlet Tessellations in the Plane”

[67]: Anglada (1997), “An improved incremental algorithm for constructing restricted Delaunay triangulations”

[68]: Si (2015), “TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator”

[69]: Shepherd et al. (2008), “Hexahedral mesh generation constraints”

[70]: Gregson et al. (2011), “All-Hex Mesh Generation via Volumetric PolyCube Deformation”

the challenges and approaches to address them, the different methods are credited.

The challenges faced are:

- 1. Geometry description** Whether it is an embedded object, a void, or a crack, its geometry must somehow be described. Different approaches to this problem include analytical descriptions of embedded domains, level-set functions or the CAD representations of the embedded objects themselves.
- 2. Integration methods** The integration of elements *cut* by the immersed boundaries requires special care. Custom integration rules must be applied to properly reflect how the discontinuity intersects with the element. Different approaches such as boundary-conform sub-integration, octree integration, or *moment-fitting* can be used.
- 3. Boundary condition application** The application of boundary conditions to cut elements, especially Dirichlet boundary conditions, is not as straightforward as in the standard finite element method and requires specific approaches, such as penalties, Lagrange multipliers or Nitsche's Method.
- 4. Stabilization methods** Numerical instability due to *small cut elements* results in unphysically steep gradients, harming both the solution itself and the performance of iterative solvers. Different approaches to stabilize the problem are the α -Method, eigenvalue-based stabilization, or the Ghost Penalty method.
- 5. Material-Material-Interfaces and Enrichment** If the immersed boundary separates two materials, jumps in the unknowns and/or their derivatives appear. Special enrichment functions can be used to enrich the space of approximation functions, beyond what standard shape functions can offer, to capture the effect of the discontinuity on the unknowns and/or their derivatives.

The following section outlines the state of the art and popular methods to address the presented challenges. One method is then selected as a suitable candidate to be applied to metal forming processes and is subsequently used in Sec. 4 to simulate *Friction Stir Welding* processes. Finally, adaptive mesh refinement is further discussed in the context of embedded domain methods.

3.2 State Of The Art

3.2.1 Geometry Description and Element Classification

All unfitted methods require a geometrical description of the embedded domain or discontinuity. This section gives an overview of two different approaches: explicit *Boundary Representations* (BREPs) and implicit level-set functions.

A classification of elements is required to identify which elements are *inside* the embedded object, which ones are *outside*, and which ones

are *cut*, as shown in Fig. 3.3. When a crack is formed, the classification requires marking those elements effected by the crack and those which harbor the crack tip. The classification is later used to assign the appropriate material properties to all elements and a customized integration rule to the cut elements. A description of the geometry is required to carry out this classification.

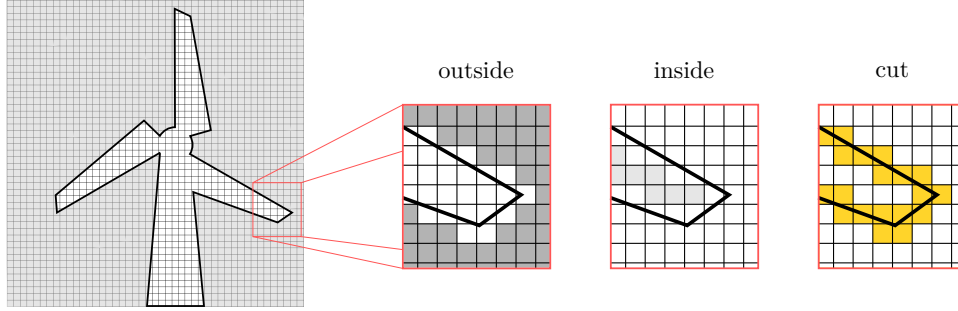


Figure 3.3: Element Classification

Explicit Geometry Description The BREP approach provides an explicit and precise description of the embedded domain's shape. This is typically achieved by importing geometry data in formats such as STL, STEP, or IGES (see Figs. 3.4a, 3.4b). The finite element algorithm can then directly access the imported geometry and check each element to determine if it intersects with the BREP [71]. In practice, this means testing all elements against the facets of the utilized STL (or NURBS surfaces of an STEP or IGES) geometry to determine which elements are cut and how. Because this approach directly utilizes the original geometry of the embedded domain, it achieves exceptional accuracy.

[71]: Meßmer et al. (2022), “Efficient CAD-integrated isogeometric analysis of trimmed solids”

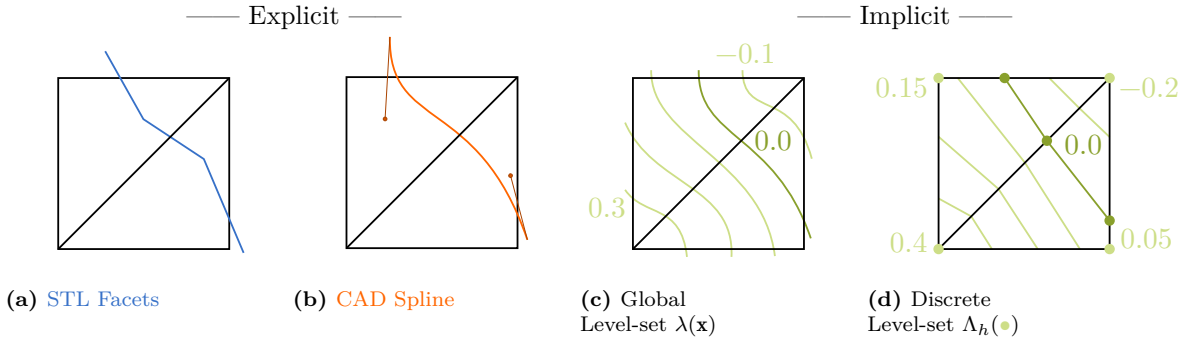


Figure 3.4: Selection of domain description methods

However, the explicit nature of this method introduces significant implementation challenges. The embedded domain's geometric complexity often translates directly into equally complex sub-integration cells. Thus, while the BREP approach offers unparalleled accuracy, it demands sophisticated algorithms to handle these complexities and only delivers fully if the integration rule is based on the same explicit geometry description.

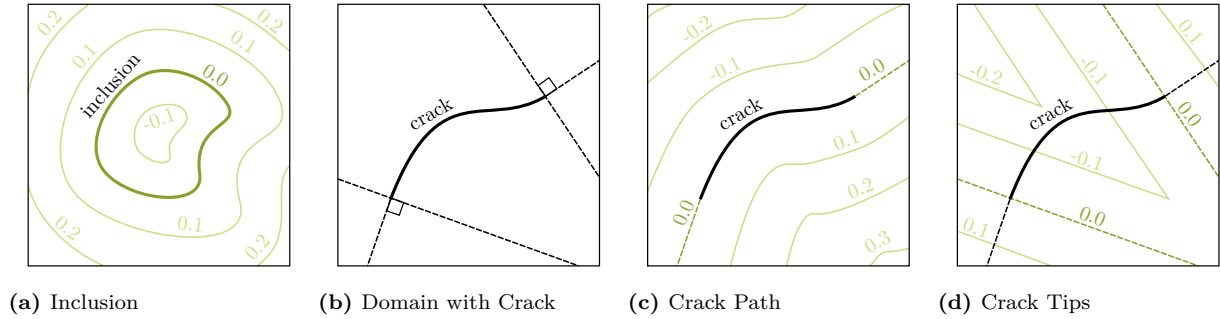
Implicit Geometry Description In contrast, implicit methods provide a more robust way to describe the domain boundary without

[72]: Sukumar et al. (2001), “Modeling holes and inclusions by level sets in the extended finite-element method”

[73]: Gibou et al. (2018), “A review of level-set methods and some recent applications”

explicitly defining its geometry. The most common implicit representation is the level-set function $\lambda(\mathbf{x})$ [72, 73]. A level-set function is an implicit geometry description because it makes no direct statement about where the boundary is but where the points of the domain lay in relation to the boundary – typically in terms of their distance from the boundary. The boundary itself is then implicitly represented by the isosurface (e.g., $\lambda(\mathbf{x}) = 0$) of the level-set function.

A simple example of a level-set function is the signed distance function, where the value at each point corresponds to the shortest (orthogonal) distance to the domain boundary. Positive values indicate points outside the domain; negative values indicate points inside, and a value of zero denotes a point on the boundary. The same approach can be used to describe a crack within a domain. One level-set function indicates the crack path, and a second one indicates the location of the crack tip.



Level-set functions can be defined globally across the entire computational domain (Fig. 3.4c) or discretely at specific points in space (Fig. 3.4d). On the one hand, global level-set functions are very precise, but the complexity of the geometry they describe is rather limited since a global continuous function must be found to fully describe the embedded domain or crack.

On the other hand, discrete level-set functions are computed by interpolating discrete level-set values Λ_i , defined at nodes $i \in I$ within each finite element using the element's shape functions $\mathbf{N}$, used to interpolate unknowns and spatial coordinates in the isoparametric finite element concept. Thus, the level-set function within an element is given by

$$\lambda_h(\mathbf{x}) = \sum_{i \in I} N_i(\mathbf{x}) \Lambda_i.$$

and is therefore defined in the same space as the finite element approximation, typically L_2 .

Since the boundary description is based on interpolation using standard shape functions, the resulting sub-integration cells have relatively simple topologies, especially when considering linear elements. Therefore, the discrete level-set method is a pragmatic choice for applications requiring reliable and efficient domain representation.

Using discrete level-set values opens up the possibility of working with CAD files without the challenges of directly handling the CAD geometry. The discrete signed distance values of the finite element nodes can be pre-computed using 3D modeling software, particularly CAD

Figure 3.5: Use of Level-set Functions for the Description of an Inclusion (a) and a Crack (c-d)

kernels such as *OpenCascade* [74]. While less accurate than the BREP method, this approach offers greater robustness and simplicity and allows for a completely automated description of the embedded domain.

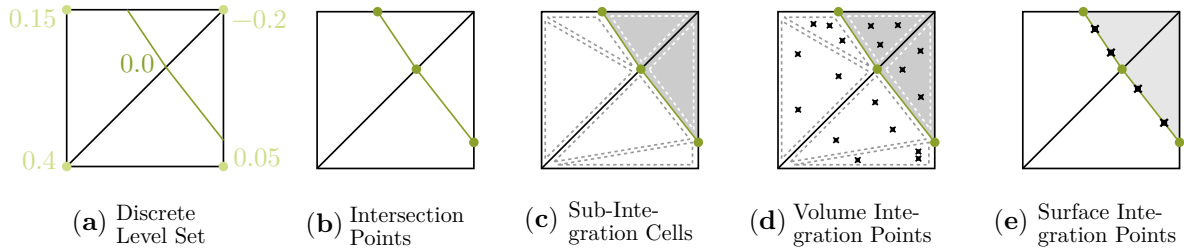
[74]: (2024), *OpenCascade*

Remark 3.2.1. *The terminology for functions used to compute nodal signed distance values may vary depending on the chosen library. However, the essential operations typically required include calculating the shortest orthogonal distance from a point to a surface and conducting point membership classification (PMC) to ascertain whether a given point is located inside or outside a closed surface. The PMC process provides the sign for the signed distance calculation.*

3.2.2 Integration Method

The second challenge when dealing with unfitted meshes is the integration of cut elements. This involves not only the integration of the (potentially discontinuous) volume but also the integration on the immersed boundary, when imposing boundary conditions. Two distinct sets of integration points are necessary: one for the volume and another for the boundary. Various techniques can be employed to determine the appropriate number, placement, and weights of these integration points. Depending on the type of interface (material-void or material-material) one or two sets of volume integration points are required.

Boundary-Conform Integration The boundary-conform integration is typically based on a discrete level-set description of the embedded domain (Fig. 3.6a) and is most often applied for linear elements (though not exclusively, see Rem. 3.2.2), especially for triangles and tetrahedra. The intersection points are computed on the element boundaries (Fig. 3.6b). Next, each resulting sub-element is subdivided into sub-integration cells of known topology (*i.e.*, triangles 2D, see Fig. 3.6c, tetrahedra in 3D), and finally Gauss-Legendre integration points are placed in the sub-integration cells. The weights of the volume integration rule are adjusted according to the ratio between the size of the sub-integration cells and the original element size.

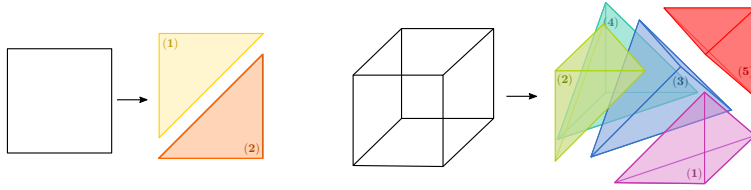


Similarly, integration points can be placed on the cut to obtain a surface integration rule to apply boundary conditions (Fig. 3.6e). Note that shape functions evaluated at these “surface” integration points contribute to all element nodes, as the integration points are located *inside* the element.

Figure 3.6: Integration Rule for boundary-conform integration of two triangle elements.

The benefits of the boundary-conform integration approach are its low computation cost and the relatively low number of integration points. However, this benefit only holds if the application is restricted to linear simplicial elements (triangles and tetrahedra). As soon as non-linear shape functions are permitted, the resulting cut is no longer a straight line, and more than two intersection points might appear. This drastically complicates the algorithmic construction of boundary-conform sub-integration cells.

However, for linear quadrilaterals and hexahedra a possible workaround exists. A decomposition into simplicial shapes can be performed. Next the previously described integration rule is then applied to each simplicial [75, 76]. A quadrilateral can be decomposed into two triangles and a hexahedron can be decomposed into, *e.g.*, five tetrahedra, as shown in Fig. 3.7.



[75]: Cheng et al. (2009), “Higher-order XFEM for curved strong and weak discontinuities”
 [76]: Kudela et al. (2015), “Efficient and accurate numerical quadrature for immersed boundary methods”

Figure 3.7: Decomposition of a quadrilateral into two triangles and a hexahedron into five tetrahedra.

Unfortunately, the creation of the integration rule is ambiguous because the orientation of the decomposition products (triangles and tetrahedra) influences how the level-set function is approximated. Imagine a quadrilateral where the discrete level-set values of adjacent nodes have alternating signs. Depending, how the quadrilateral is decomposed (bottom left to top right $\square \rightarrow \nearrow$, as in Fig. 3.7, or top left to bottom right $\square \rightarrow \searrow$), the resulting iso-surface of the interpolated level-set function and its normal surface will look differently. However, by reducing the mesh size, the impact of this ambiguity diminishes.

Quadtree/Octree Integration This second approach to defining an integration rule for cut elements aims to subsequently refine the integration cell based on a space-tree algorithm, *i.e.*, octree, or quadtree. After each level of refinement, the newly created sub-integration cells are tested to determine whether they are cut by the boundary. The cut sub-integration cells are then further refined until the desired accuracy is reached.

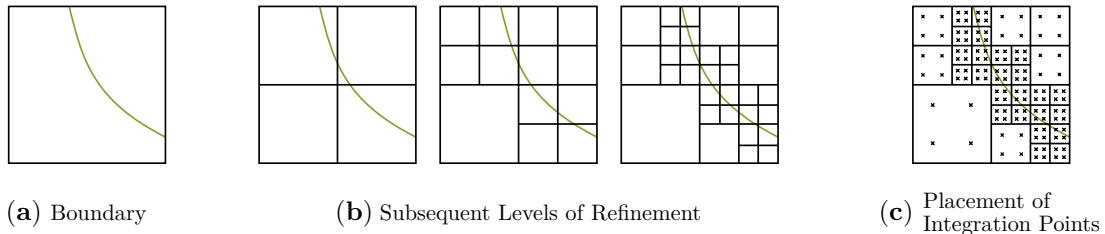


Figure 3.8: Octree refinement of sub-integration cells.

Quadtree/Octree integration is a very robust method that works well with arbitrarily high- or low-orders shape functions and different methods of boundary description. However, it generally results in a large

number of integration points, as is already evident from Fig. 3.8c. Different approaches were proposed to reduce the final number of integration points by either merging sub-integration cells [77] or hierarchically overlaying differently fine sub-integration meshes and performing boolean operations on their outcome [78].

This method of integration is generally only used for volume integrals. A parametric description of the boundary needs to be constructed to integrate boundary terms. One option is to represent the boundary with a triangulation (either constructed or taken from an explicit boundary description such as an STL file) and integrate the triangulation using numerical integration methods [79].

Remark 3.2.2. *The differences between apparently distinct approaches to integrate cut elements, are rather blurry. One example is the Smart Octree method by Kudela et al. [80]. See also reference [81]. Like the quadtree/octree procedure, the cut element is subdivided into four or eight parts. The newly created nodes are moved onto the interface, allowing for a more accurate and boundary-conform integration. Most importantly, this also applies to higher-order elements, resulting in higher-order, boundary-conform sub-integration cells. Therefore, the methods can be seen both as a modification of the quadtree/octree approach and as an extension of the previously described boundary-conform integration to higher-order elements.*

Moment Fitting Boundary-conform integration is easy to implement for linear elements. It yields a relatively low number of integration points but is cumbersome to implement for higher-order elements. On the other hand, space-tree integrations are very robust and work well regardless of the order of shape functions and complexity of the immersed boundary. They, however, yield much higher numbers of integration points.

Moment fitting [81, 82] is a method that aims unify the benefits of both methods: a general algorithm that applies regardless of the order of shape functions yet yields a moderate number of integration points and is applicable to define both volume and surface integration rules. The only limitation is that the geometry must be described using a level-set function. Though, as introduced above, an embedded CAD geometry can be easily converted into a discrete level-set function using a CAD library.

The idea of moment fitting is to construct a custom integration rule (*i.e.*, location and weights of integration points) for each cut element based on how it is cut. The custom numerical integration rule is manufactured to give the same result as an analytical integration over the same domain

$$\sum_{i=1}^N f_j(\mathbf{x}_i)w_i = \int_{\Omega} f_j(\mathbf{x}) dV, \quad j = 1, \dots, M \quad (3.1)$$

where $\mathbf{x}_i$ are the positions of the N integration points, w_i are the corresponding weights and f_j are a set of M integrands, also called *moments*.

[77]: Pető et al. (2022), “Octree-based integration scheme with merged sub-cells for the finite cell method: Application to non-linear problems in 3D”

[78]: Pető et al. (2023), “Boolean finite cell method for multi-material problems including local enrichment of the Ansatz space”

[79]: Düster et al. (2008), “The finite cell method for three-dimensional problems of solid mechanics”

[80]: Kudela et al. (2016), “Smart octrees: Accurately integrating discontinuous functions in 3D”

[81]: Hubrich et al. (2017), “Numerical integration of discontinuous functions: moment fitting and smart octree”

[81]: Hubrich et al. (2017), “Numerical integration of discontinuous functions: moment fitting and smart octree”

[82]: Müller et al. (2013), “Highly accurate surface and volume integration on implicit domains by means of moment-fitting”

The integration rule can be obtained by solving the moment-fitting equations

$$\begin{bmatrix} f_1(\mathbf{x}_1) & \cdots & f_1(\mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ f_M(\mathbf{x}_1) & \cdots & f_M(\mathbf{x}_N) \end{bmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = \begin{pmatrix} \int_{\Omega} f_1 dV \\ \vdots \\ \int_{\Omega} f_M dV \end{pmatrix} \quad (3.2)$$

$$A_{ji}w_i = b_j \quad (3.3)$$

$$\mathbf{A}\mathbf{w} = \mathbf{b}. \quad (3.4)$$

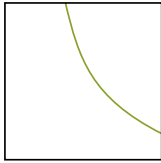
The system is linear with respect to the weights but non-linear concerning the positions of the integration points. For this reason, solving for both the weights and positions is computationally expensive. It is only used when the level-set function does not change over time. A cheaper approach is to adapt to a set of integration points (typically standard Gauss-Legendre integration points) and solely determine the weights. The right-hand side of Eq. 3.2 cannot be evaluated as is, however, through a series of transformations, the volume integral $\int_{\Omega} f_j(\mathbf{x}) dV$ can be replaced by a number of surface integrals over its boundary. The same hierarchical replacement is performed until the original integral is expressed by a set of line integrals, evaluated using numerical integration. A surface integration rule is constructed analogously.

Typically, the number of nodes exceeds the number of moments $N > M$, resulting in an under-determined system. The resulting optimization problem

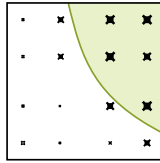
$$\min w_i w_i \text{ s.t. Eq. (3.3)} \quad (3.5)$$

is solved using a least squares approach.

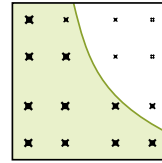
Ultimately, three integration rules are obtained: two volume integration rules (one for either side of the cut), and a surface integration rule. The curiosity of this method is that all three integration rules might share the same integration point locations, and none of them are located on the cut, nor exclusively on the side of the cut that is sub-integrated.



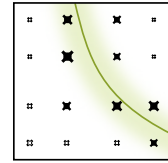
(a) Boundary



(b) Volume Int. Rule 1



(c) Volume Int. Rule 2



(d) Surface Int. Rule

Remark 3.2.3. *If not further constrained, Eq. (3.5) permits negative solutions, which results in negative weights. This in itself is not a problem. However, negative weights can harm the convergence of many iterative solvers. Instead, a non-negative least squares solver can be used to ensure only positive weights are obtained [83].*

Figure 3.9: Volume and Surface Integration Rules obtained with Moment Fitting.

[83]: Garhuom et al. (2022), “Non-negative moment fitting quadrature for cut finite elements and cells undergoing large deformations”

3.2.3 Boundary Condition Application

To explain the different methods of applying boundary conditions in cut elements, it is useful to recall the weak form of the mixed u/p problem shown in Eq. (2.75).

$$\begin{aligned} & B^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p]) + B_{\text{oss}}^{\text{up}}(\delta p, p) \\ & - \int_{\Gamma_D} \underline{\delta \mathbf{u} \cdot \mathbb{C}' : \nabla^s \mathbf{u} \cdot \mathbf{n} \, dA} - \int_{\Gamma_D} \underline{\delta \mathbf{u} \cdot \mathbf{n} \, p \, dA} \\ & - \int_{\Gamma_D} \underline{\delta p \mathbf{u} \cdot \mathbf{n} \, dA} = L(\delta \mathbf{u}) \end{aligned} \quad (3.6)$$

The additional terms do not appear in the standard finite element method since they are solely defined on the Dirichlet boundary ($\mathbf{u} = \bar{\mathbf{u}}$ on Γ_D), where the test functions $\delta \mathbf{u}$ and δp vanish if the Dirichlet boundary conditions are prescribed strongly at the boundary nodes. However, this is no longer true if the boundary conditions are applied weakly. This might be the case when the boundary doesn't adhere to the (non-conforming) mesh and the weak imposition of Dirichlet boundary conditions becomes the only option.

Remark 3.2.4. *On the contrary, the application of Neumann boundary conditions ($\boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}}$ on Γ_N) is straight forward and is accomplished by integrating the prescribed tractions on the surface integration points that were defined in the previous section.*

Penalty Method The easiest method to apply the Dirichlet boundary condition is to add an additional term that enforces $\mathbf{u} = \bar{\mathbf{u}}$ on Γ_D using a penalty, α [84, 85]. The problem then reads

$$\begin{aligned} & B^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p]) + B_{\text{oss}}^{\text{up}}(\delta p, p) + \underline{B_{\Gamma_D}^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p])} \\ & + \int_{\Gamma_D} \delta \mathbf{u} \cdot \alpha (\mathbf{u} - \bar{\mathbf{u}}) = L(\delta \mathbf{u}), \end{aligned} \quad (3.7)$$

where B_{Γ_D} refers to the bilinear form of the additional terms introduced in Eq. (3.6). The drawback of this method is that the correct value of the penalty α is unknown and depends both on the problem (e.g., Young's modulus or viscosity) and the mesh size h . A penalty that is too low does not fully enforce the boundary condition and a penalty that is too large harms the conditioning and can cause oscillations or locking.

Nitsche's Method To remedy this problem, a study of coercivity is conducted to determine a suitable value for α [86]. In the context of linear elasticity, this amounts to

$$\alpha > 2k C_{\text{inv}}/h, \quad (3.8)$$

where k is a measure related to the material stiffness, e.g., Young's modulus E , $C_{\text{inv}} > 0$ is a constant to be chosen by the user and

[84]: Babuska (1973), "The Finite Element Method with Penalty"

[85]: Babuška et al. (1973), "Nonconforming Elements in the Finite Element Method with Penalty"

[86]: Nitsche (1971), "Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind"

[87]: Ern et al. (2004), *Theory and Practice of Finite Elements*

originates from the application of the inverse inequality [87]. It depends on the elements shape and polynomial degree; h is the element size.

The problem is also symmetrized and ultimately reads

$$B^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p]) + B_{\text{oss}}^{\text{up}}(\delta p, p) + \underline{B_{\Gamma_D}^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p])} \\ - \int_{\Gamma_D} \underline{\mathbb{C}' : \nabla^s \delta \mathbf{u} \cdot \mathbf{n} \cdot (\mathbf{u} - \bar{\mathbf{u}})} \, dA + \int_{\Gamma_D} \delta \mathbf{u} \cdot \alpha(\mathbf{u} - \bar{\mathbf{u}}) = L(\delta \mathbf{u}). \quad (3.9)$$

Lagrange Multiplier An alternative method of applying Dirichlet boundary conditions in weak form is adding an additional equation to enforce the boundary condition and an additional unknown, called a Lagrange multiplier $\boldsymbol{\lambda}$ – together with its corresponding test function $\boldsymbol{\mu}$ [88].

$$B^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p]) + B_{\text{oss}}^{\text{up}}(\delta p, p) + \underline{B_{\Gamma_D}^{\text{up}}([\delta \mathbf{u}, \delta p], [\mathbf{u}, p])} = L(\delta \mathbf{u}) \\ \int_{\Gamma_D} \delta \mathbf{u} \cdot \boldsymbol{\lambda} \, dA + \int_{\Gamma_D} \boldsymbol{\mu} \cdot (\mathbf{u} - \bar{\mathbf{u}}) \, dA = 0 \quad (3.10)$$

The resulting system closely resembles the mixed u/p problem, where the pressure effectively acts as a Lagrange multiplier enforcing the incompressibility constraint $\nabla \cdot \mathbf{u} = 0$. As in the mixed u/p formulation, the system takes the form of a saddle-point problem, which is inherently unstable for arbitrary combinations of approximation function spaces and may therefore require stabilization.

3.2.4 Small Cut Stabilization

Interfaces that separate a physical domain from a non-physical one (*e.g.*, voids) are a common occurrence in embedded domain methods. When dealing with such an interface, a choice must be made regarding the integration strategy of cut elements. One option is to integrate over the entire element, encompassing both physical and non-physical parts. While this approach enhances numerical robustness, it compromises consistency, as the integration over the non-physical parts can influence the solution within the physical domain. Alternatively, integration can be restricted solely to the physical parts. This approach preserves accuracy, but the condition number of the finite element matrix now depends on how the mesh is cut. Small cuts can lead to ill-conditioned matrices and threaten the stability of the discrete scheme [89]. This issue becomes particularly critical when imposing Dirichlet boundary conditions or performing domain decomposition on unfitted meshes, where the *inf-sup* condition must be satisfied, as is the case for the mixed problems presented in Secs. 2.3.1 and 2.3.2 [90].

α -Method This first approach, also known as the *material stabilization technique* [79, 91], addresses the problem by applying a soft ma-

[88]: Glowinski et al. (1999), “A distributed Lagrange multiplier/fictitious domain method for particulate flows”

[89]: Burman (2010), “Ghost penalty”

[90]: Burman et al. (2014), “CutFEM: Discretizing geometry and partial differential equations”

[79]: Düster et al. (2008), “The finite cell method for three-dimensional problems of solid mechanics”

[91]: Dauge et al. (2015), “Theoretical and Numerical Investigation of the Finite Cell Method”

terial to the non-physical domain, thus stabilizing the problem. The application is realized by multiplying the material properties with

$$\alpha(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_{\text{phy}}, \\ 10^{-q}, & \text{otherwise.} \end{cases} \quad (3.11)$$

According to Garhuom, Usman, and Düster [92], the q value should be chosen in a range of $q = 5, \dots, 9$ to achieve reasonable results without losing too much accuracy. The major advantage of this method is its simple implementation since it only relies on a few additional integration points in the non-physical sub-integration cell to achieve stability.

Eigenvalue Stabilization The underlying problem of small-cut instabilities are close-to-zero eigenvalues of the corresponding (element) stiffness matrix. These close-to-zero eigenvalues cause linear dependencies, ultimately leading to an ill-conditioned problem. However not all deformation modes are equally affected. By factorizing the element stiffness matrix into an eigenvalue and -vector decomposition, it becomes possible to selectively stabilize the affected eigenmodes [92–94].

The stabilized stiffness matrix reads

$$\mathbf{K}_{\text{stab}} = \mathbf{K} + \sum_{i=1}^n \epsilon_i \tilde{\mathbf{v}}_i \otimes \tilde{\mathbf{v}}_i, \quad (3.12)$$

where $\mathbf{K}$ is the original stiffness matrix, $\tilde{\mathbf{v}}_i$ are the eigenvectors ($i = 1, \dots, n$) of near-zero eigenmodes of $\mathbf{K}$ and ϵ_i is a stabilization factor that needs to be chosen. Alternatively, the stabilization factor can be defined as a function of the corresponding near-zero eigenvalue and a set of constants $\epsilon_i = f(\lambda_i, c)$.

Remark 3.2.5. *Note that every element stiffness matrix naturally has a number of zero eigenvalues due to rigid body motions. The eigenvalue stabilization method described above does not address this issue but only close-to-zero eigenvalues caused by the small cuts.*

Ghost Penalty The idea of the Ghost Penalty (GP) approach is to apply a penalty in the interface zone that extends the coercivity of the physical domain. Without the GP, coercivity is only satisfied in the physical domain; hence, it is also satisfied only in the physical part of the cut elements. The GP approach extends the coercivity also to the non-physical parts of the cut elements [89]. Typically the penalty is applied to the cut elements and the first two layers of physical elements, as shown in Fig. 3.10. In the following this selection of elements is denoted as $K(\Omega^{1\text{st}2})$.

This increases the robustness of the problem without sacrificing accuracy. In particular, it removes the dependence of the condition number on the boundary orientation with respect to the mesh. The condition number instead scales the same as for the standard (*i.e.*, body fitted) finite element method.

[92]: Garhuom et al. (2022), “An eigenvalue stabilization technique to increase the robustness of the finite cell method for finite strain problems”

[92]: Garhuom et al. (2022), “An eigenvalue stabilization technique to increase the robustness of the finite cell method for finite strain problems”

[93]: Loehnert (2014), “A stabilization technique for the regularization of nearly singular extended finite elements”

[94]: Loehnert (2015), “Stabilizing the XFEM for static and dynamic crack simulations”

[89]: Burman (2010), “Ghost penalty”

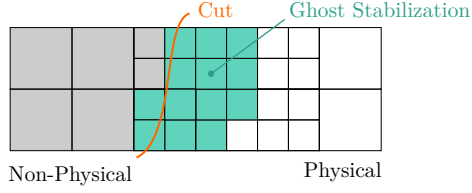


Figure 3.10: Cut elements and two layers of physical elements, stabilized using the Ghost Penalty method.

The conditioning is enhanced by stabilizing the gradients of the unknowns (*i.e.*, $\nabla \mathbf{u}$ and ∇p) using a penalty term. In the original method, the penalization is applied to the jumps of the gradient across the element faces (or edges in 2D), though an equivalent method of performing the penalization exists utilizing orthogonal projections,

$$\begin{aligned}
 & B^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) + B_{\text{oss}}^{\text{up}}(\delta p_h, p_h) \\
 & \quad + B_{\Gamma_D}^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) \\
 & \quad + \sum_K \gamma_u \int_{K(\Omega^{1^{\text{st}}_2})} \nabla \delta \mathbf{u}_h \cdot \mathbf{P}_h^\perp[\nabla \mathbf{u}_h] \, dV \\
 & \quad + \sum_K \gamma_p \int_{K(\Omega^{1^{\text{st}}_2})} \nabla \delta p_h \cdot \mathbf{P}_h^\perp[\nabla p_h - \mathbf{b}] \, dV = L(\delta \mathbf{u}_h),
 \end{aligned} \tag{3.13}$$

with the stabilization parameters

$$\gamma_u = \chi_u 2G \quad \text{and} \quad \gamma_p = \chi_p \frac{h^2}{2G}, \tag{3.14}$$

where χ_u and χ_p are (user defined) stabilization constants.

Remark 3.2.6. *The Ghost Penalty stabilization terms resemble the OSS stabilization terms introduced in Eq. 2.72. Both include projections of the gradients of the unknown, orthogonal to the finite element approximation space. Both are applied on a by-element basis. However, the Ghost Penalty terms are only applied to a small number of elements (the cut elements plus the first two layers). In contrast, the OSS stabilization is applied to the entire physical domain.*

3.2.5 Enrichment

In the standard finite element method, high gradients require h - or p -refinement, and actual discontinuities require the mesh to conform to it. However, a different approach exists by *enriching* the polynomial approximation space in a way that non-smooth solutions can be modeled independently from the mesh.

In practice, this means either replacing original shape functions with enriched ones (*intrinsic* enrichment) or adding the enrichment to the original ones (*extrinsic* enrichment). Enrichment can either be global (*i.e.*, enrichment is applied at all nodes, I) or local (*e.g.*, at particular points of interest, such as cracks and crack-tips or material interfaces), restricting the enrichment to a subset of nodes I^* .

Eq. 3.15 shows the augmented approximation of the unknowns. N_i and N_i^* are standard finite element shape functions for node i , which

are often (but not necessarily) the same. $\mathbf{U}_i$ are the standard FE nodal unknowns and $\mathbf{A}_i$ is a set of additional nodal unknowns.

$$\mathbf{u}_h(\mathbf{x}) = \underbrace{\sum_{i \in I} \mathbf{N}_i(\mathbf{x}) \mathbf{U}_i}_{\text{std. FE approx.}} + \underbrace{\sum_{i \in I^*} \mathbf{N}_i^*(\mathbf{x}) \cdot [\psi(\mathbf{x}) - \psi(\mathbf{x}_i)] \mathbf{A}_i}_{\text{enrichment}} \quad (3.15)$$

An important aspect of the enrichment terms is that they (same as in the standard FEM) need to respect the partition of unity, meaning

$$\sum_{i \in I} \mathbf{N}_i(\mathbf{x}) = \sum_{i \in I^*} \mathbf{N}_i^*(\mathbf{x}) = 1. \quad (3.16)$$

The resulting contour of the unknown is qualitatively shown in Fig. 3.11.

The function $\psi(\mathbf{x})$ is the actual enrichment function that holds the special information about the kind of high gradient or discontinuity present in this part of the computational domain. For example, by introducing an absolute value function, weak discontinuities (*i.e.*, kinks, see Fig. 3.11, top) can be approximated and by introducing a step function, such as the Heaviside function, strong discontinuities are introduced (see Fig. 3.11, bottom).

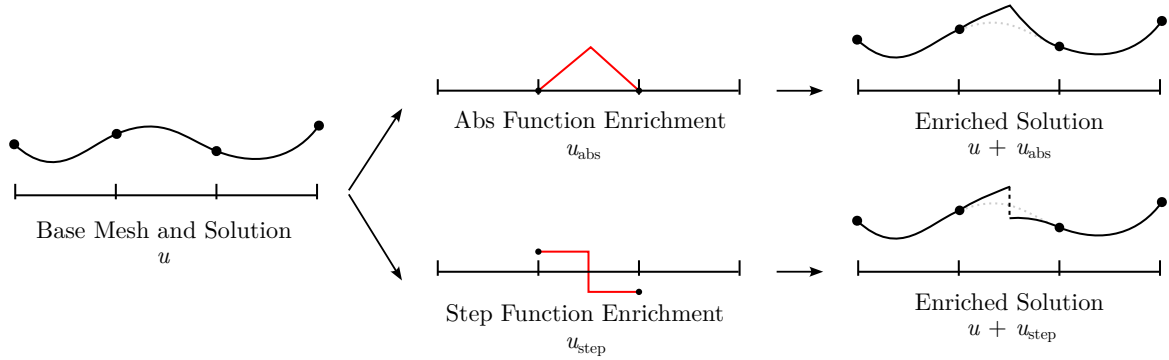


Figure 3.11: Local enrichment of approximation space

3.2.6 Established Methods

Over the past 30 years, a multitude of methods that allow for unfitted meshes have been proposed. As mentioned initially, a clear distinction is not always possible.

The Extended Finite Element Method (XFEM) [95, 96], Generalized Finite Element Method (GFEM) [97, 98] and Partition of Unity Method (PUM) [99–101] were all developed with the idea in mind to deal with crack formations and material discontinuities. Typically, the physical domain is described using a body-fitted mesh, and only the discontinuities do not conform to the mesh. Over time, the distinction between these methods has become less clear, and Fries and Belytschko [102] argue that they are almost identical today. These methods typically deal with material-material interfaces, hence enrichment is an integral part of these methods.

Immersed Boundary Methods (IBM) [103, 104], Fictitious Domain Methods (FDM) [105, 106], the Finite Cell Method (FCM) [79, 107] and CutFEM [108–110] form the other big group of methods. They typically focus more on embedding inclusions or voids in physical domains or the entire physical domain in a fictitious one with the aim of using a regular mesh. As a result, the interfaces are often material-void interfaces (though not exclusively). Thus, enrichment is often optional since the solution in the non-physical domain is of no interest. However, enrichment can be used also with these methods (*e.g.*, local enrichment in FCM [78]).

Today, the most used methods are the Finite Cell Method and CutFEM. FCM is particularly popular in the engineering community and typically uses higher-order shape functions. As a result, boundary-conform integration is rather complicated, and the methods of choice are either quadtree/octree integration or Moment Fitting. Furthermore, the preferred methods of stabilization are the α -method and eigenvalue stabilization. Enrichment is used when necessary. The Spectral Cell Method (SCM) [111] is very closely related to the Finite Cell Method and resembles the “unfitted version” of the Spectral Element Method (SEM) [112].

The latter method (CutFEM) has a very strong mathematical background and enjoys great popularity in the fluid-dynamics community (although not exclusively), where linear elements are preferred. Ghost Penalty stabilization is strongly linked with this approach, and thanks to linear (or low-order) elements, boundary-conform integration of cut elements is the norm. (Adaptive) mesh refinement is often used to better capture the embedded boundary. Again, enrichment is used when necessary.

Regarding the application of boundary conditions, no specific preference exists. However, one notable unfitted approach is the Shifted Boundary Method (SBM) [113, 114], which was specifically developed to strongly enforce Dirichlet boundary conditions on non-mesh-aligned interfaces by projecting them onto nearby nodes.

The description of the embedded domain or interface is typically chosen independently from the employed method. However, implicit geom-

[95]: Moës et al. (1999), “A finite element method for crack growth without remeshing”

[96]: Sukumar et al. (2000), “Extended finite element method for three-dimensional crack modelling”

[97]: Strouboulis et al. (2000), “The design and analysis of the Generalized Finite Element Method”

[98]: Strouboulis et al. (2000), “The generalized finite element method: an example of its implementation and illustration of its performance”

[99]: Melenk et al. (1996), “The partition of unity finite element method: Basic theory and applications”

[100]: Babuška et al. (1997), “The Partition of Unity Method”

[101]: Griebel et al. (2000), “A Particle-Partition of Unity Method for the Solution of Elliptic, Parabolic, and Hyperbolic PDEs”

[102]: Fries et al. (2010), “The extended/generalized finite element method: An overview of the method and its applications”

[103]: Roma et al. (1999), “An adaptive version of the Immersed Boundary Method”

[104]: Peskin (2002), “The immersed boundary method”

[105]: Hansbo et al. (2002), “An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems”

[106]: Ramière et al. (2007), “A fictitious domain approach with spread interface for elliptic problems with general boundary conditions”

[79]: Düster et al. (2008), “The finite cell method for three-dimensional problems of solid mechanics”

[107]: Parvizian et al. (2007), “Finite cell method”

[108]: Burman et al. (2010), “Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method”

[109]: Burman et al. (2012), “Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method”

[110]: Burman et al. (2014), “Fictitious domain methods using cut elements: III. A stabilized Nitsche method for Stokes’ problem”

[78]: Petö et al. (2023), “Boolean finite cell method for multi-material problems including local enrichment of the Ansatz space”

[111]: Duczek et al. (2014), “Numerical analysis of Lamb waves using the finite and spectral cell methods: LAMB WAVE ANALYSIS USING THE FCM AND THE SCM”

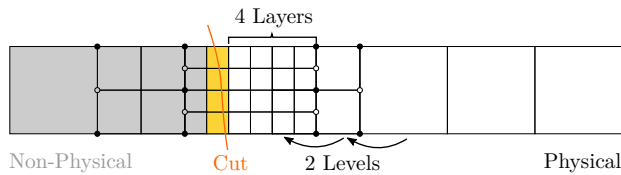
etry description based on level-set functions (especially in a discretized manner) is very popular because the interface is expressed using the same shape functions as the solution of the problem. Furthermore, it is easy to generate a discretized level-set function from an explicit CAD geometry.

3.3 Mesh Refinement and Multi-Refinement Criteria

An accurate representation of the immersed boundary is crucial for both geometric fidelity and solution accuracy. In the Finite Cell Method (FCM), high-order shape functions are typically employed, ensuring a precise geometry description, when using a discrete level-set function. Alternatively the geometry could be described using an explicit representation or a global level-set function. All these options yield high geometric accuracy. As a result, the integration method must also be sufficiently precise to fully exploit this advantage; otherwise, the benefits of an accurate geometry representation are lost. In contrast, CutFEM typically assumes linear elements and a discrete level-set function. In this case *Boundary-Conform Integration*, as described in Sec. 3.2.2, already yields the most accurate integration possible. However, the geometry description is less accurate due to the approximation with linear shape functions. As a result, in CutFEM, the mesh needs to be refined, not the integration scheme.

Adaptive mesh refinement (as briefly mentioned in Sec. 2.5.3), in particular quadtree/octree refinement, is a good tool to address this issue. Similar to the hierarchical subdivision of integration cells of the quadtree/octree integration scheme, a parent element marked for refinement is split into four or eight child elements. Like any other adaptive refinement scheme, the quadtree/octree method needs to be driven by a refinement criterion. In the context of unfitted methods, the focus lies on approximating the immersed boundary as precisely as possible to reduce the geometrical approximation error. However, the area close to the immersed boundary is usually of significant interest, and adequate discretization is required even at some distance from the immersed boundary.

A suitable refinement criterion is the *Layers-and-Levels* approach [115], designed to achieve a specified number of subdivisions for cut elements (levels) and a defined number of the small elements adjacent to the interface (layers), see Fig. 3.12.



[112]: Patera (1984), “A spectral element method for fluid dynamics: Laminar flow in a channel expansion”

[113]: Main et al. (2018), “The shifted boundary method for embedded domain computations. Part I: Poisson and Stokes problems”

[114]: Main et al. (2018), “The shifted boundary method for embedded domain computations. Part II: Linear advection–diffusion and incompressible Navier–Stokes equations”

[115]: Baiges et al. (2017), “An adaptive Fixed-Mesh ALE method for free surface flows”

Figure 3.12: Refinement levels and layers

Additionally a *two-one balancing* condition ensures that the refinement levels of neighboring elements differ by no more than one, limiting the number of hanging nodes.

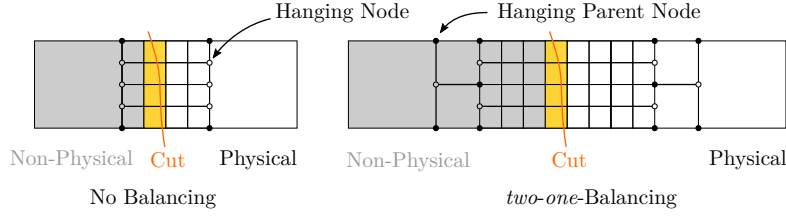


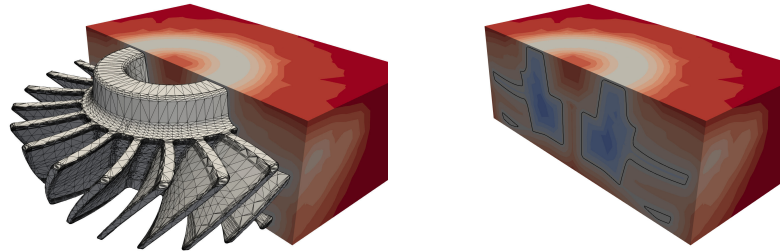
Figure 3.13: Balancing of refinement levels

If hanging nodes are not constrained, the result is a non-conform interpolation of the unknowns. To remedy this, the shape functions can be modified to become conforming [115, 116]. The hanging nodes are eliminated from the finite element equations, and the contribution of their test and shape functions is added to the corresponding hanging parent node (hp). The shape functions are updated as follows,

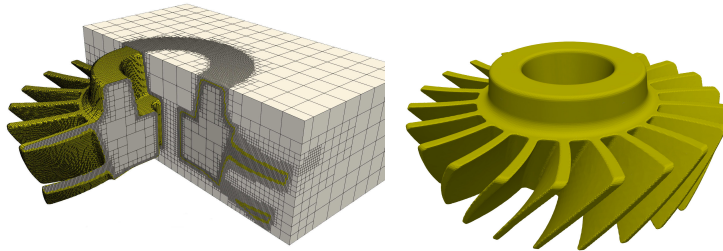
$$\mathbf{N}_{hp}^*(\mathbf{x}) = \mathbf{N}_{hp}(\mathbf{x}) + \sum_{i=1}^{n_{child}} \mathbf{N}_{np}(\mathbf{x}_i) \mathbf{N}_{hci}(\mathbf{x}), \quad (3.17)$$

where $\mathbf{N}_{hp}$ and $\mathbf{N}_{hci}$ denote the original shape and test functions for the parent and children nodes respectively, $\mathbf{N}_{hp}^*(\mathbf{x})$ denotes the modified hanging parent shape function, and $\mathbf{x}_i$ denotes the position in space of the i^{th} hanging child node. The values of the unknowns at the hanging nodes are obtained by interpolation from nodes of the parent element.

Fig. 3.14 depicts the full workflow from CAD geometry (here STL) to obtain the computational domain of the physical in the Finite Element framework. Octree mesh refinement is used together with the levels and layers refinement criterion.



(a) STL Geometry Embedded into Domain (b) Level-Set Function with Iso-Zero Surface



(c) Mesh and Element Classification (d) Physical Domain

Figure 3.14: Workflow of Embedding CAD Geometry into Computational Domain, Computing Level-Set Function and Classifying Elements

[115]: Baiges et al. (2017), “An adaptive Fixed-Mesh ALE method for free surface flows”

[116]: Baiges et al. (2017), “Reficientlib: An Efficient Load-Rebalanced Adaptive Mesh Refinement Algorithm for High-Performance Computational Physics Meshes”

3.4 Concluding Remarks

The final step is to determine which unfitted finite element method is most suitable to be used for metal forming and manufacturing applications. This decision must be made taking into account the methods presented in Sec. 2 (*Element Technology*) to deal with locking, isochoric behavior and strain and stress accuracy.

For the given problem, the best choice is the CutFEM approach, which means using linear elements, a discrete level-set function based on a CAD geometry and the Ghost Penalty approach to stabilize small cut elements. Adaptive refinement based on the layers-and-levels refinement criterion is used to enhance the accuracy of the geometry description. The reasoning leading to this choice is as follows.

Sec. 2.3 (*Mixed Formulations*) introduced the mixed two-field u/p and mixed three-field $u/p/e$ formulations as very robust methods to tackle fully isochoric problems, eradicate volumetric locking and ensure local convergence when using linear elements. Unfortunately the resulting problems are not *inf-sup* stable. Sec. 2.4 (*Numerical Stability and Stabilization*) introduced the Variational Multiscale Method (in particular Orthogonal subscales, OSS) to stabilize the problems. However OSS requires projecting the approximated subscales onto a space orthogonal to the space of finite element approximations. The most accurate method of resolving the projections, which does not add yet more unknowns, is to evaluate the projections using the previous iteration's solution, *i.e.*, a fixed-point iteration. This effectively introduces “a non-linearity” to the problem. However, in the context of metal forming and manufacturing applications, this hardly matters since kinematic and strong material non-linearities are present anyway. The introduced non-linearity from the orthogonal projections simply adds to the existing non-linearities.

Although their origin is entirely different, the Ghost Penalty terms bear a strong resemblance with the OSS stabilization terms since they also rely on said orthogonal projections, which facilitates implementation. Furthermore the non-linearity caused by the Ghost Penalty projections also simply adds to the existing non-linearity. In both cases, linear elements allow for the use of lumped mass matrix projections without risking negative mass matrix entries. Ultimately the same considerations regarding the projections that are required for OSS apply for the Ghost Penalty method.

This means the application of the Ghost Stabilization is relatively easy and blends in well with the OSS stabilization. Furthermore the use of linear elements highly benefits the OSS stabilization approach since this reduces the number of stabilization terms required to satisfy the *inf-sup* conditions.

The need to describe geometry using a discrete level-set function, a requirement in the CutFEM approach, poses no significant challenge, as this function can be efficiently computed from the CAD geometry, as previously discussed. In fact, this approach is more robust compared to directly working with the CAD file, making it a more reliable choice for complex geometries.

All things considered, the CutFEM approach is the most suitable method for metal forming and manufacturing applications. Sec. 4 (*Friction Stir Welding*) covers different approaches to simulating Friction Stir Welding and argues that the use of unfitted meshes can open up new possibilities.

The question of how to apply the boundary conditions is deliberately left unanswered at this point and is picked up in Sec. 4.3.3 (*Boundary Conditions*), since it must be viewed in conjunction with the modeling of FSW problems. Further mentions of the *stabilized u/p formulation on an unfitted mesh* refer to the following problem:

$$\begin{aligned}
& \int_{\Omega} \nabla^s \delta \mathbf{u}_h : \mathbb{C}' : \nabla^s \mathbf{u}_h \, dV + \int_{\Omega} \nabla \cdot \delta \mathbf{u}_h p_h \, dV \\
& + \int_{\Omega} \delta p_h \nabla \cdot \mathbf{u}_h \, dV - \int_{\Omega} \delta p_h p_h / \kappa \, dV \\
& - \sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV \\
& - \int_{\Gamma_D} \delta \mathbf{u}_h \cdot \mathbb{C}' : \nabla^s \mathbf{u}_h \cdot \mathbf{n} \, dA \\
& - \int_{\Gamma_D} \delta \mathbf{u}_h \cdot \mathbf{n} p_h \, dA - \int_{\Gamma_D} \delta p_h \mathbf{u}_h \cdot \mathbf{n} \, dA \\
& + \sum_K \gamma_{1K} \int_{K(\Omega^{1st2})} \nabla \delta \mathbf{u}_h \cdot \mathbf{P}_h^\perp [\nabla \mathbf{u}_h] \, dV \\
& + \sum_K \gamma_{2K} \int_{K(\Omega^{1st2})} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h - \mathbf{b}] \, dV \\
& + [\text{Boundary Condition Terms}] = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_N} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA.
\end{aligned} \tag{3.18}$$

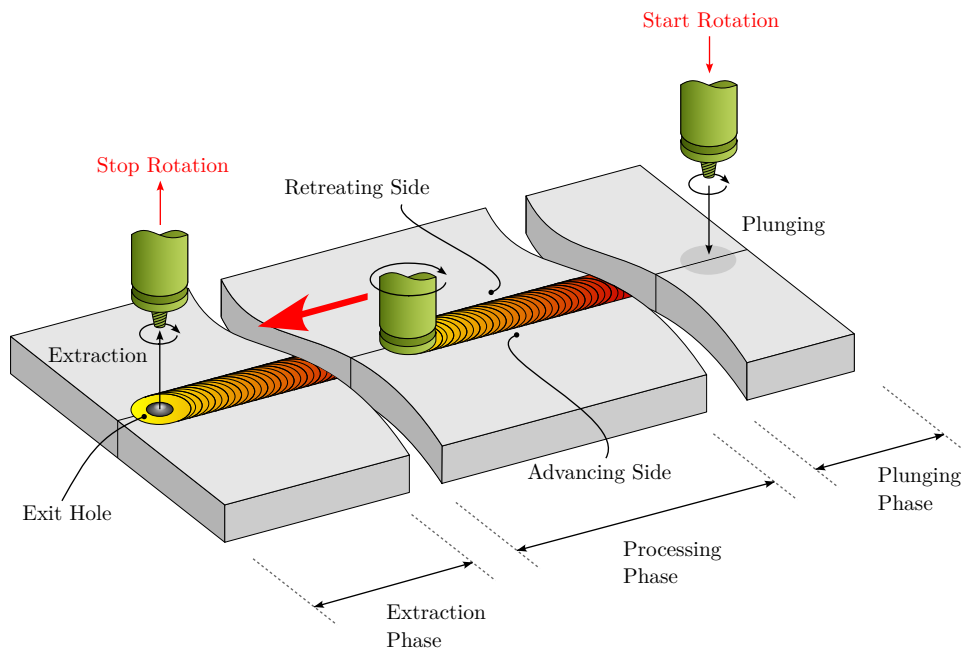
$$\begin{aligned}
& B^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) + B_{\text{oss}}^{\text{up}}(\delta p_h, p_h) \\
& + B_{\Gamma_D}^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) + B_{\text{gs}}^{\text{up}}([\delta \mathbf{u}_h, \delta p_h], [\mathbf{u}_h, p_h]) = L(\delta \mathbf{u}_h)
\end{aligned} \tag{3.19}$$

The third part of this work addresses Friction Stir Welding (FSW) as an emerging solid-state joining technique and different approaches to simulate the process. FSW is a very complex and sensitive process characterized by a strong thermo-mechanical coupling, temperature dependent material properties, isochoric material flow, very high strains and strain rates, and complex geometries and boundary conditions. The previous sections about *Element Technology* (Sec.2) and *Unfitted FEM / Embedded Domain Methods* (Sec.3) laid the groundwork to address this complex manufacturing process and develop a suitable numerical framework to simulating it.

However, before developing the numerical framework, the process itself and its state of the art are described. Furthermore, the state of the art of FSW simulations is presented to justify why the previously presented advanced finite element technologies are a valuable addition.

4.1 State of the Art FSW Technology

Friction Stir Welding is a solid-state joining process in which two facing workpieces are welded together using a rotating, non-consumable pin-tool that is first plunged into the materials, moved along their interface, and extracted after completing the weld, see Fig.4.1.



The joining process is achieved because of the softening of the material due to heating caused by plastic dissipation of the workpiece materials and friction between tool and workpiece. Typically, the material is not melted, which constitutes the first big advantage of FSW over conversational (*i.e.*, light-beam based) welding techniques. The process

Figure 4.1: Friction Stir Welding Process

was developed at *The Welding Institute* (TWI) in Cambridge, UK, in 1991 and has been patented by Thomas et al. [117]. Following, the process is explained in greater detail, and advantages and challenges are presented to motivate the necessity of numerical simulations to advance the process and help engineers produce high-quality welds.

Remark 4.1.1. *There are several related processes that are not the focus of this work. However, certain aspects of these processes, such as heat generation or microstructure evolutions, are also relevant to FSW. Therefore, these related processes are occasionally mentioned. Examples include Rotary Friction Welding (RFW), Linear Friction Welding (LFW), Friction Stir Spot Welding (FSSW), and Friction Stir Processing (FSP).*

Remark 4.1.2. *A substantial portion of FSW research is experimental. Consequently, studies frequently concentrate on the impact of one parameter at a time, occasionally two or three in more extensive studies. However, the combinations of tool shapes, process parameters, weld dimensions, and materials (including dissimilar pairings) are practically limitless. Due to the process's highly interconnected and non-linear nature, it is challenging to generalize findings. This section provides a fundamental overview of various aspects of FSW, such as the influence of tool shapes, process parameters, applicability to different materials, and the advantages over conventional welding techniques. The aim is not to review specific experimental findings, which is why this section primarily references review publications.*

Welding Process The to-be-welded workpieces are placed on a solid, supporting surface (*backplate*), butted together, and rigidly clamped to suppress any movement. The non-consumable pin-tool is rotated at a constant speed and plunged into the welding line between the two workpieces [118]. When coming into contact, the material is heated up due to the occurring friction, is initially displaced, and finally stirred in place. The tool remains rotating with a constant speed but otherwise remains stationary for some amount of time (*tool dwell time*) for the temperature to reach a sufficient level [119]. The primary cause of the heating is friction between the tool's shoulder, which rests on the workpiece's surface. The (featured) pin of the tool is responsible for mixing the material. Through a combination of heating and stirring, the material close to the tool is plastified at very high strain rates [118]. This concludes the *plunging phase*.

In the *welding or processing phase*, the still rotating tool moves along the weld line, continuously heating and stirring the material. The material is moved by the pin, aided by the features of the tool (see Par. *Tool Shape*), forced around the tool to its backside and fills the hole in the tool's wake [118]. After cooling, a solid and continuous joint between the two workpieces is obtained with the characteristic surface state, depicted in Fig. 4.2.

It is important to note that FSW is *not* a symmetric process. Due to the simultaneous rotation and advancement of the tool, the material flow differs on both sides. The side where the direction of the tool's rotational velocity aligns with the advancing direction is commonly

[117]: Thomas et al. (1995), *FRIC-TION WELDING*

[118]: Dialami (2014), "Thermo-mechanical analysis of welding processes"

[119]: Kilic et al. (2023), "A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties"

[118]: Dialami (2014), "Thermo-mechanical analysis of welding processes"

[118]: Dialami (2014), "Thermo-mechanical analysis of welding processes"

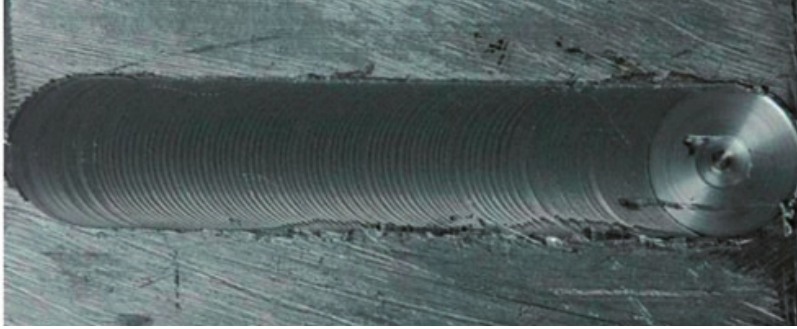


Figure 4.2: Completed, sound weld.

Figure origin: [120] (cropped)

referred to as the *advancing side*. The other side, where both velocity components are opposed, is typically called the *retracting side*.

Affected Zones The welded region is classified into distinct zones based on the thermal and mechanical effects induced by the rotating and translating tool, see Bastier et al. [121] and Mishra et al. [122].

The Stir Zone (SZ), also known as the Weld Nugget (WN), is the area directly affected by the tool pin. Due to the frictional and stirring action of the tool, it undergoes intense plastic deformation and experiences high temperatures. This leads to dynamic recrystallization, resulting in the formation of fine, equiaxed grains that are significantly smaller than those in the base material.

Adjacent to the SZ is the Thermo-Mechanically Affected Zone (TMAZ), which is mechanically and thermally influenced. The deformation in this region occurs due to the forging action of the tool shoulder, causing the grains to align along the welding direction. The grain size in the TMAZ is larger than in the SZ but is still affected by the welding process.

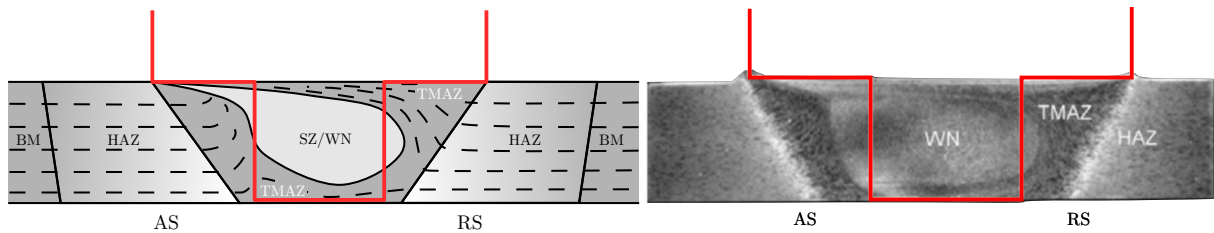


Figure 4.3: Different zones obtained after FSW

Figure origin: [123] (cropped, modified, relabeled)

Beyond the TMAZ lies the Heat-Affected Zone (HAZ), which is exposed to elevated temperatures but experiences no plastic deformation. While the mechanical properties of this region are altered due to the generated heat, the metallurgical changes resemble those observed in traditional fusion welding processes. The grain size in the HAZ is larger than in the TMAZ.

Finally, the Base Material (BM) or Parent Material Zone remains unaffected by the welding process, retaining its original mechanical and metallurgical properties.

Machines Friction Stir Welding (FSW) can be performed on both conventional milling machines and CNC machines (Fig. 4.4a), provided that an appropriate tool and suitable process parameters are used [119].

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

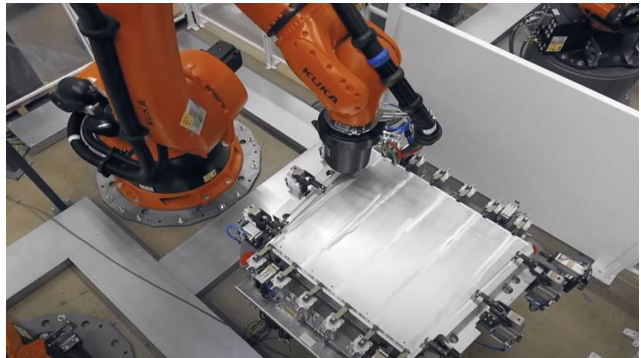
However, specialized machines are typically preferred for large-scale operations. The most commonly used systems include gantry-based systems (Fig. 4.4b), which offer higher rigidity and minimal deflection, and robotic arm systems (Fig. 4.4c), which provide greater versatility but are more prone to deflection.



(a) Milling machine, modified for friction stir welding ^[a]



(b) ESAB SuperStir™ FSW equipment ^[b]



(c) KUKA KR FORTEC 500 R2830 MT ^[c]

To ensure process stability and weld quality, modern FSW installations integrate advanced sensing and control systems, including real-time temperature distribution monitoring, force feedback, and adaptive control algorithms that enable precise process adjustments [119].

Tool Shape The tool (Fig. 4.5) is the central part of the welding setup. Its shape and dimensions have a huge impact on the resulting weld quality. The tool serves three primary functions [118]:

Heating of the workpiece During the welding process a vertical force or prescribed plunging depth presses the shoulder against the top surface of the workpiece. Due to this pressure and the relative movement of the shoulder and workpiece, friction occurs. The friction is one of the two sources of heating.

Figure 4.4: FSW machines

^a <https://bil-ibs.be/...>
(accessed: 12.03.2025)

^b <https://t-welding.com/...>
(accessed: 12.03.2025)

^c <https://www.kuka.com/...>
(accessed: 12.03.2025)

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

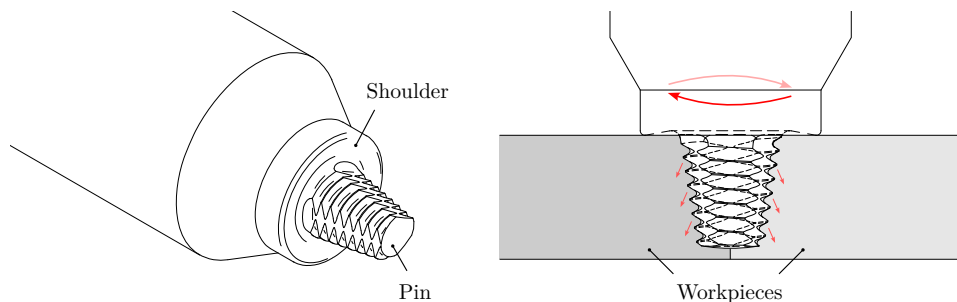
[118]: Dialami (2014), “Thermo-mechanical analysis of welding processes”

Containment The shoulder is also responsible for containing the material flow during the welding process and preventing it from being forced out of the weld.

Stirring material The pin is responsible for stirring the material and ultimately producing the joint. This motion is heavily influenced by the shape of the pin (dimensions and distinctive features like threads, flats, or flutes). The energy dissipated due to the plastification of the material is also converted to heat and is the second source of heating.

The tool is typically manufactured from hard and durable materials such as tungsten carbide, hot work tool steels, Ni-Co alloys, or refractory materials. It must withstand high compressive loads, resist creep, fatigue, and various forms of wear at different temperatures, maintain fracture toughness, be machinable for fabrication, and remain chemically stable with both the workpiece and environment [124]. The thickness of the to-be-welded material dictates its overall dimensions. Most importantly, the length of the pin has to be slightly shorter than the thickness of workpieces to ensure that the material is properly joined over the entire thickness [118, 119].

The shape of the tool has a huge impact on the characteristics and quality of the weld, and many different shapes and features have been proposed in recent times. This brief overview only covers classical tools (as shown in Fig. 4.5), which consist of *one* shoulder and *one* pin, and which are permanently fixed together.



[124]: Maji et al. (2022), “An overview on friction stir welding/processing tools”

[118]: Dialami (2014), “Thermo-mechanical analysis of welding processes”

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

Figure 4.5: Pin-tool

Remark 4.1.3. *There are a number of new inventions, such as adjustable tools, where the pin can be retracted after completing the weld – filling out the exit-hole [125], and the so-called Bobbin tool [126], where a second shoulder is pressed against the bottom of the workpieces instead of a rigid backing plate.*

Within the group of classical tools, different aspects can vary. The shoulder end surface can be flat, concave, or convex and contain features such as scrolls, ridges, knurling, grooves, or concentric circles (see Fig. 4.6). The pin tip can be flat or domed, and the outer surface’s general shape can be cylindrical or tapered. Regardless of the general shape, the cross-section profile of the pin can either be circular or polygonal; the pin can be flatted or contain flutes, and it can be threaded or not (see Fig. 4.7).

[125]: Ding (1998), “Retractable Pin-Tool Technology for Friction Stir Welding”

[126]: Marie et al. (2005), “Development of the Bobbin Tool Technique on Various Al-Alloys”

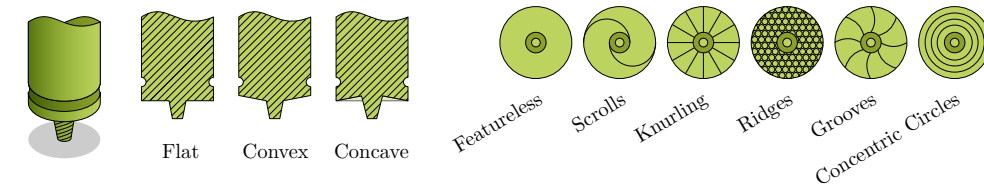


Figure 4.6: Shoulder Types

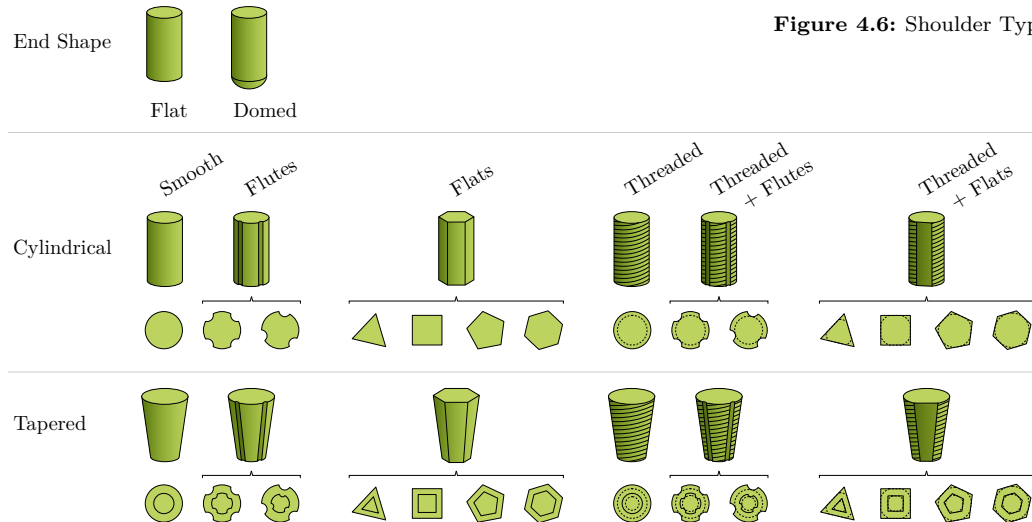


Figure 4.7: Pin Types

Different features affect the material flow, plastification, heat generation, surface finish, affinity to defect formation, reaction forces, torque, and many more. It is impossible to give a comprehensive explanation of all relations since the number of possible combinations is too large. The effects vary based on the tool size (in conjunction with the material thickness), the to-be-welded *Materials* and the *Process Parameters*.

However, a few findings can be stated regardless. Flat shoulder end surfaces can lead to excessive flash. Concave shoulders are a good solution to addressing this problem, as the displaced material is contained by the concave shoulder cavity. Furthermore, shoulder features like scrolls or grooves facilitate material movement from the shoulder edge towards the pin, further preventing flash [124].

Fluted or flattened pins increase material flow by trapping material and partially acting like a cutter, increasing plastification and causing finer grains. Threaded pins further increase material flow and facilitate pore elimination and breaking of the oxide layer. Tools with left-handed threads are rotated clockwise and vice versa [124]. This motion “unscrews” the tool, pushing down the material (see Fig. 4.5, right).

However, featured pins are subject to higher wear than featureless pins and might not be suitable for highly abrasive or high-strength alloys [124]. The ratio between pin length and shoulder diameter is important to obtain a defect-free weld and depends on the plate thickness and conductivity of the material. Studies have also shown that the torque and advancing forces increase with a decreasing number of flats [124].

Process Parameters The process parameters of FSW are

— *rotation speed* or *tool rotation rate*,

[124]: Maji et al. (2022), “An overview on friction stir welding/processing tools”

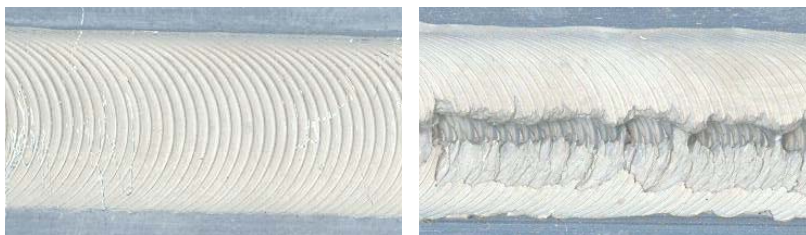
- *tool travel rate, advancing speed or welding speed,*
- *plunging depth* or – as a different means of a pushing the tool into the workpiece – *axial force or plunging force,* and
- *tool tilt angle.*

The rotation speed strongly influences the total heat [118], where larger rotational speeds cause higher temperatures [119] due to increased friction and plastic deformation in the material. Furthermore, the rotation speed affects the size of the stirred area, the microstructure, and ultimately, the tensile strength of the welds [127]. Too low rotation speeds can lead to a significant reduction in micro hardness in the SZ [127] and too high rotation speeds can cause the creation of tunnel defects [119].

The advancing speed also influences the total heat [118], though in an opposite manner. Higher advancing speeds lead to lower temperatures along the welding line [119], because the tool remains less time at the same positions, giving the material less time to heat up. Too high advancing speeds have been associated with the formation of wormhole defects [119].

A tilt angle is introduced to influence the contact pressure under the shoulder. The typical angles range between 1° and 3°. The tool is tilted backward (with respect to the welding direction), causing the rear edge of the shoulder to produce deeper into the material while lowering the pressure under the leading edge of the shoulder or even causing incomplete contact [118]. Tool tilt helps prevent internal void formation, such as tunnel defects [128], and reduces flashing.

FSW systems typically operate under either force control or displacement control. In force control, the machine maintains a constant axial force throughout the welding process, ensuring stable pressure at the weld interface. This enhances material flow and weld quality, making it particularly beneficial for welding materials with varying thicknesses or properties. In displacement control, the tool's plunging depth is prescribed, ensuring consistent penetration into the material. According to ESAB [129], force-controlled FSW processes are always to be preferred over displacement-controlled processes due to the possible appearance of welding defects, as shown in Fig. 4.8.



[118]: Dialami (2014), “Thermo-mechanical analysis of welding processes”

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

[127]: Kosturek et al. (2022), “The influence of tool rotation speed on the low-cycle fatigue behavior of AA2519-T62 friction stir welded butt joints”

[128]: AluStir (—), *FSW: Tool Tilt - How far Should the Tool be Tilted during Friction Stir Welding?*

[129]: ESAB (2004), *Friction Stir Welding - the ESAB Way*

Figure 4.8: Possible weld defects due to displacement controlled setup (right), compared to force controlled process (left).

Temperatures The peak temperatures typically range between 80% and 95% of the melting temperature of the workpiece material [119]. This means that while the material is significantly softened to facilitate the stirring, it is not melted. These lower temperatures (compared to, e.g., arc welding) minimize the HAZ, which leads to more durable welds [119]. Due to the unsymmetric nature of the process, the temperatures

on the advancing side are slightly higher than on the retreating side [130].

The temperature evolution has a large impact on the precipitates dissolution and re-precipitation, thus influencing the welds hardness.

Materials Over the past decades FSW has been widely used for joining aluminum alloys [118], which are especially difficult to weld using traditional welding techniques such as arc welding. Aluminum rapidly oxidizes when exposed to the atmosphere, and the resulting oxide layer has a much higher melting point than the unoxidized material. Furthermore, molten aluminum absorbs high amounts of hydrogen, weakening the joint, and, as a remedy, shielding gas might be required. Furthermore, hot cracking can occur due to high levels of thermal stress. The combination of being a (comparatively) soft material, which allows for tools manufactured from regular tool steel, and being especially affected by the phase change occurring in traditional welding, makes aluminum a perfect candidate for FSW.

Magnesium is another (comparatively) soft material, which is also known to be difficult to weld for similar reasons as aluminum. This includes high oxidation tendency, high thermal conductivity, low melting point, and susceptibility to hot cracking. FSW has been used to weld magnesium alloy components [131] with great success and yields a number of benefits over conventional welding, such as improved microstructure, absence of cracking, non-requirement of shielding gas, and reduced energy consumption [132].

Copper and its alloys (*e.g.*, bronzes, brasses, and beryllium copper) are important industrial materials due to their high thermal and electrical conductivity, good mechanical and corrosion resistance, as well as ductility [133]. Unfortunately, the otherwise advantageous property of high thermal conductivity causes difficulties in welding. To maintain a temperature above the melting point, more heat must be input compared to materials with lower thermal conductivity. This can lead to increased oxidation and a generally larger HAZ [134, 135]. In FSW, the heat is generated very locally, and the material is not melted. Hence, it is also a good choice for copper alloys and has been used in this way [136–138].

Steels are undoubtedly one of the most important construction materials, and traditional steel welding techniques are well established. However, the phase change can also cause problems in traditional steel welding, although the problem is not as pronounced as for other materials. Nevertheless, the advantages of FSW can also be capitalized for steel welding. Similar to the above examples, the main advantages are the reduced HAZ size and excellent mechanical and corrosion resistance properties of the welds [139]. The main obstacles for steel FSW are the increased material properties of the base material, the resulting tool damage, and insufficient heat generation. To overcome these difficulties, different in-situ heating solutions have been proposed, such as laser or induction heating [140].

Titanium is typically well weldable using conventional techniques. However, workpiece distortion and generally poor weld quality can be an is-

[130]: Verma et al. (2017), “Study on temperature distribution during Friction Stir Welding of 6082 aluminum alloy”

[118]: Dialami (2014), “Thermo-mechanical analysis of welding processes”

[131]: Desai et al. (2021), “Friction stir welding of AZ31 magnesium alloy: A review”

[132]: Singh et al. (2018), “Review on friction stir welding of magnesium alloys”

[133]: Bulacu et al. (2024), “A Review on Friction Stir Welding of Copper: Tool Geometry, Process Parameters, and Joint Properties”

[134]: Mishra et al. (2005), “Friction stir welding and processing”

[135]: Nandan et al. (2008), “Recent advances in friction-stir welding – Process, weldment structure and properties”

[136]: Çam (2011), “Friction stir welded structural materials: beyond Al-alloys”

[137]: Nielsen (2012), “Modeling and Control of Friction Stir Welding in 5 cm thick Copper Canisters”

[138]: Heidarzadeh et al. (2020), “Development of a fuzzy logic based model to elucidate the effect of FSW parameters on the ultimate tensile strength and elongation of pure copper joints”

[139]: Mohan et al. (2021), “A Review on Friction Stir Welding of Steels”

[140]: Álvarez et al. (2013), “Assisted Friction Stir Welding of Carbon Steel: Use of Induction and Laser as Preheating Techniques”

sue. Other possible problems include overheating, excess grain growth, and high residual stresses due to low heat conductivity [141, 142]. The main problems faced in titanium FSW are the fast wear of the tool and the intermixing of wear particles into the weld [142].

Dissimilar material joints are particularly challenging and complicated to achieve using conventional welding techniques. The main problems are the different melting points of the materials (leading to uneven flow in the weld zone) and different thermal expansion coefficients (leading to residual stresses and weld deformations). Since joining materials below their melting points, FSW can prevent these problems and has been proven as suitable for creating dissimilar aluminum-to-aluminum joints [143] and even aluminum-to-non-aluminum joints [144].

Joint Properties The main properties to which FSW joints are optimized are tensile strength, ductility, hardness, fatigue strength, and residual stresses.

The tensile strength of load-bearing structures is very important, and research regarding the optimization of FSW joints in this regard is an ongoing topic. The strength of the joint is affected by many factors, including process parameters and tool design [119]. Typically, FSW welds perform very well in this regard, reaching weld efficiencies* of 75% to 100% [119]. The high joint strength can be attributed to the smaller grain sizes in the dynamically recrystallized zone (SZ/WN, [145]).

The change in ductility of an FSW joint, compared to the BM, is not uniform and varies depending on the different zones: HAZ, TMAZ, and WN/SZ (see [134]). Mahoney et al. [146] studied the mechanical properties of FSW joints in 7075Al-T651 aluminum, both in the longitudinal direction (*i.e.*, within the weld nugget) and transverse to the weld direction. Within the WN, elongation did not change significantly compared to the BM. However, for the specimen cut transverse to the weld direction (thus containing microstructures from all four zones), elongation was reduced by almost 50%. A study by Sato and Kokawa [147] on the transverse tensile properties of friction stir welds in 6063-T5 aluminum also showed a similar reduction of 50%.

Hardness is typically linked to the size and distribution of precipitates. The FSW process involves temperature-related precipitate dissolution and re-precipitation, which largely depend on the obtained temperatures and cooling rates, which, in turn, depend on the welding speed [145].

Cyclic loading can cause FSW-manufactured components to fail due to fatigue cracking. As is typical for fatigue, the failure can occur at stresses much smaller than the static strength limit and without significant deformation prior to the failure. According to Li, Gao, and Li [148], when welding with optimized process parameters (*i.e.*, when no internal defects or flaws are present), the fatigue crack initiation generally starts at the weld surface, and typically between TMAZ and HAZ. Apart from defects, the fatigue performance of FSW joints is mainly affected by process parameters, stress ratio, and residual stresses.

[141]: Junaid et al. (2017), "Comparison of microstructure, mechanical properties, and residual stresses in tungsten inert gas, laser, and electron beam welding of Ti-5Al-2.5Sn titanium alloy"

[142]: Tarasov et al. (2022), "Friction Stir Welding of Ti-6Al-4V Using a Liquid-Cooled Nickel Superalloy Tool"

[143]: Sen et al. (2020), "Investigations into FSW joints of dissimilar aluminum alloys"

[144]: Shankar et al. (2022), "Dissimilar friction stir welding of Al to non-Al metallic materials: An overview"

[119]: Kilic et al. (2023), "A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties"

[145]: Kossakowski et al. (2019), "Effect of Selected Friction Stir Welding Parameters on Mechanical Properties of Joints"

[134]: Mishra et al. (2005), "Friction stir welding and processing"

[146]: Mahoney et al. (1998), "Properties of friction-stir-welded 7075 T651 aluminum"

[147]: Sato et al. (2001), "Distribution of tensile property and microstructure in friction stir weld of 6063 aluminum"

[145]: Kossakowski et al. (2019), "Effect of Selected Friction Stir Welding Parameters on Mechanical Properties of Joints"

[148]: Li et al. (2018), "Fatigue of Friction Stir Welded Aluminum Alloy Joints: A Review"

* *Weld efficiency* is defined as the ratio between joint strength and BM strength

Although friction stir welded components are subjected to lower peak temperatures compared to arc-welded ones, the total power input is, in fact, similar. In fusion welding, despite the higher peak temperatures, the molten material cannot carry any load. As a result, the residual stresses in FSW and arc welding are quite similar [149]. As in conventional welding, the residual stresses can cause performance degradation and distortions. Furthermore, it is very well known that residual stresses negatively impact the fatigue strength of materials [150]. The manipulation of process parameters only provides limited control over the resulting residual stresses. However, traditional techniques to reduce residual stresses post-weld, such as peening and, especially, rolling, have proven to be effective [149].

Advantages FSW has several advantages over conventional welding techniques – many of which have been mentioned when FSW’s applicability to different *Materials* was discussed. To summarize: FSW allows for welding a wide range of materials (or dissimilar material pairs), some of which have been viewed as *un-weldable* with traditional methods. Many of these difficult-to-weld materials are characterized by a high susceptibility to oxidation, together with the formation of oxide layers with a significantly higher melting point, which are strong thermal conductors. In both cases, conventional welding requires a higher heat input, leading to an enlarged HAZ and other associated problems. In FSW, the overall heat input is lower, mitigating some of the aforementioned problems.

FSW does not melt the to-be-welded workpiece, entirely avoiding the formation of Hot Cracking phenomena such as Solidification Cracking or Liquation Cracking, as these are directly linked to phase change. For the same reason, problems associated with porosity and degradation are avoided [119].

Further advantages are the excellent joint properties, such as high strength, fatigue resistance, and toughness, and relatively fast process speed, which allows for efficient production of components [119]. These properties can be largely attributed to the dynamic recrystallization in the stir zone and the excellent microstructure obtained in the final weld.

Finally, FSW is a sustainable manufacturing process for multiple reasons [151]. First, the environmental impact and possible health implications are minimized due to the absence of shielding gases, fluxes, filler materials, and electrode coatings, all of which can generate hazardous fumes. Second, FSW is an energy-efficient process, consuming less energy than gas metal arc welding for comparable welds. Shrivastava, Kronen, and Pfefferkorn [152] showed an energy reduction by 40% for aluminum 6061-T6 plates and Dawood, Mohammed, and Rajab [153] showed a fourfold reduction for aluminum 1030 plates.

Applications FSW is actively used in a variety of engineering applications. Although there is still a lot of research to be done, and numerous *Challenges* to be faced, FSW is an established manufacturing technique for certain applications.

[149]: Williams et al. (2010), “Residual stresses in friction stir welding”

[150]: Almen et al. (1963), *Residual Stresses and Fatigue in Metals*

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

[151]: Majeed et al. (2021), “Friction stir welding: A sustainable manufacturing process”

[152]: Shrivastava et al. (2015), “Comparison of energy consumption and environmental impact of friction stir welding and gas metal arc welding for aluminum”

[153]: Dawood et al. (2014), “Advantages of the Green Solid State FSW over the Conventional GMAW Process”



(a) Falcon 9 booster tank made of aluminum lithium alloy with FSW joints ^[a]



(b) Welding underway on Orion for first artemis mission landing astronauts on the moon ^[b]



(c) FSW train panel of Bombardier's trains for London Underground ^[c]



(d) Extruded aluminum panels joined with FSW ^[d]

In spacecraft engineering [154], FSW has been successfully used to weld fuel tanks since the late 1990s and early 2000s. This includes the Delta II, IV and Atlas V rockets. More recent examples are the NASA Space Launch System (SLS) Core Stage, Falcon 9 booster tanks (Fig. 4.9a) and the Orion spacecraft (4.9b). A non-US example for the use of FSW in spacecraft manufacturing is the H-IIB launch vehicle jointly developed by Japan Aerospace Exploration Agency and Mitsubishi Heavy Industries, Ltd. [155].

In rail transportation systems, FSW is used to manufacture structural components made of aluminum. Examples include roof panels for Alstom LHB trains for the Danish State Railways manufactured at Marine Aluminium under a contract with Sapa (now Hydro), DB (German Railway) Class 423 Trains, side Panels for the carriages of London Underground (Bombardier) and car bodies (walls, roof and floor) of the A-Train by Hitachi made of aluminum extrusions that are joined using FSW [156–158]. A case study was conducted to test if components of the Medellín metro can be manufactured using FSW instead of traditional arc welding [159]. The part was manufactured on a conventional milling machine.

[154]: Wang et al. (2018), “Friction stir welding of high-strength aerospace aluminum alloy and application in rocket tank manufacturing”

[155]: Nimura et al. (2013), “Technologies and Prospects of the H-IIB Launch Vehicle”

Figure 4.9: Application of FSW transportation.

^a <https://en.m.wikipedia.org/...> (accessed: 06.02.2025)

^b <https://nasa.gov/...> (accessed: 06.02.2025)

^c <https://image.jimcdn.com/...> (accessed: 07.02.2025)

^d <https://image.jimcdn.com/...> (accessed: 07.02.2025)

In the automotive industry, FSW is used to manufacture different components. Examples include spot-joined rear doors and bonnet of the Mazda RX-8 [160], the center tunnel of the Fort GT [161], and wheels of the Volvo XC90 SUV [162] where a cast wheel hub is joined to a wrought rim.

FSW has also been used to manufacture consumer products such as the Apple MacBook frame [163].

Tang et al. [164] investigated FSW as an approach for fabricating and repairing components of nuclear power plants, and the Swedish Nuclear Fuel and Waste Management Company (SKB) has evaluated FSW as a method to seal the 5 cm thick copper canisters containing spent nuclear fuel [137].

Challenges FSW is a promising joining technique capable of welding various metals, including dissimilar metal pairs. However, it comes with certain challenges.

One frequently cited issue is the variation in material properties across different weld zones and the presence of residual stresses parallel to the weld. Additionally, post-weld inspections can be challenging, as non-destructive testing methods often prove inadequate [119]. However, these challenges are not unique to FSW but are also encountered in traditional welding techniques.

A distinct challenge specific to FSW is the need for specialized machinery – either dedicated FSW equipment or at least appropriately adapted milling machines – both of which are large and heavy machines. Moreover, the process requires high clamping forces to securely hold the workpieces in place. This restricts FSW applications to larger, industrial use cases, making handheld operation impossible. Furthermore, the initial cost of setting up an FSW production line is significantly higher than that of conventional welding methods.

Another major challenge is the process's sensitivity to its parameters. Tool design, advancing and rotational speed, and plunging depth and force all profoundly impact weld quality. Optimizing these factors is crucial not only for refining the weld's microstructure but also for preventing defects such as voids, lack of bonding, and tunnel formation.

This sensitivity stems from the complex thermo-mechanical coupling and the highly nonlinear material behavior of the process. Multiple interacting phenomena co-occur, making it nearly impossible to isolate the effect of a single process parameter or tool feature on a specific joint characteristic.

Many studies have attempted to experimentally quantify the impact of welding parameters on weld quality and defect formation. However, these experiments are costly, and their findings are not easily transferable across different materials, workpiece thicknesses, or tool geometries. Since FSW applies to a vast range of material combinations, the number of potential variations is virtually limitless.

This combination of (1) the highly nonlinear nature of the process, (2) the narrow operating range for most parameters, and (3) the countless possibilities in terms of materials, workpiece shapes, and optimization

[156]: AluStir (—), *FSW: Rolling Stock*

[157]: TWI (2002), *Railway Rolling Stock Manufacturers Implement Friction Stir Welding*

[158]: Gay et al. (2014), “Crashworthiness of Joints in Aluminium Rail Vehicles”

[159]: Hoyos et al. (2021), “Case Study: Implementation of FSW in the Colombian Rail Transport Sector”

[160]: MachineDesign (2003), *Car bodies first to be friction stir welded*

[161]: TWI (2004), *Friction Stir Welding in Series Production*

[162]: TWI (—), *Friction Welded Wheels provide improved performance and reduced running costs*

[163]: AppleInsider (2012), *Apple slims down iMac 40% with ‘friction-stir welding’ & ditching the disc drive*

[164]: Tang et al. (2022), “Friction Stir Welding – an Advanced Approach to Repair Nuclear Power Plant Components”

[137]: Nielsen (2012), “Modeling and Control of Friction Stir Welding in 5 cm thick Copper Canisters”

[119]: Kilic et al. (2023), “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”

goals makes FSW a difficult technique to master. Numerical modeling offers a valuable solution, helping researchers deepen their understanding of the process while allowing engineers to accelerate production by reducing the time required for parameter optimization.

4.2 State of the Art FSW Simulations

Finite element simulations have proven valuable tools for simulating different aspects of FSW. Early thermo-mechanical FEM models were proposed by Chao, Qi, and Tang [165] and Chen and Kovacevic [166]. They were used to study the thermal history, stress distribution, and resulting mechanical forces.

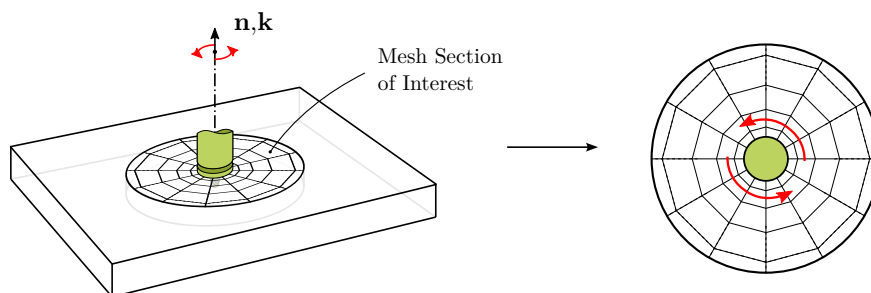
Most recently, particle methods such as Smoothed-Particle Hydrodynamics (SPH) [167, 168] and the Movable Cellular Automata Method (MCA) [169] have gained traction. Notable examples of FSW simulations using these particle methods include the works of Pan et al. [170], Jamshidi Aval [171], and Shishova et al. [172], who utilized SPH to study temperature distributions, microstructures, and bonding mechanisms, respectively. Additionally, Smolin et al. [173] employed MCA to investigate optimal ratios of advancing and rotation velocities.

Although particle methods are an emerging numerical approach, the vast majority of published studies still rely on thermo-mechanically coupled FEM simulations to predict material flow, thermal evolution, metallurgy, and potential defects. Rather than providing a broad overview of the various aspects of FSW explored through numerical methods, this section focuses specifically on the different mechanical frameworks used for thermo-mechanically coupled FSW simulations. The goal is to outline the existing frameworks, highlight their advantages and disadvantages, and clarify when the choice of framework dictates the scope of analysis versus when it remains relatively independent.

The different frameworks studied in this section are

1. *Lagrangian*,
2. *Eulerian*,
3. *Arbitrary-Lagrangian-Eulerian (ALE)*,
4. the *Apropos Kinematic Framework (AKF)* or Rigid-ALE, and
5. *Non-Body Fitted approaches*.

All frameworks lead to different kinds of mesh movement during the simulation. To illustrate the movements, a circular portion of the mesh close to the tool is studied, as shown in Fig. 4.10.



[165]: Chao et al. (2003), “Heat Transfer in Friction Stir Welding—Experimental and Numerical Studies”

[166]: Chen et al. (2004), “Thermo-mechanical modelling and force analysis of friction stir welding by the finite element method”

[167]: Monaghan (1992), “Smoothed Particle Hydrodynamics”

[168]: Liu et al. (2010), “Smoothed Particle Hydrodynamics (SPH): an Overview and Recent Developments”

[169]: Psakhie et al. (1995), “Method of movable cellular automata as a tool for simulation within the framework of mesomechanics”

[170]: Pan et al. (2013), “A new smoothed particle hydrodynamics non-Newtonian model for friction stir welding: Process modeling and simulation of microstructure evolution in a magnesium alloy”

[171]: Jamshidi Aval (2023), “Smoothed-Particle Hydrodynamics (SPH) Simulation of AA6061-AA5086 Dissimilar Friction Stir Welding”

[172]: Shishova et al. (2022), “Reversible inter-particle bonding in SPH for improved simulation of friction stir welding”

[173]: Smolin et al. (2018), “Understanding the mechanisms of friction stir welding based on computer simulation using particles”

Figure 4.10: FSW mesh detail

Due to high temperatures close to the melting point, the workpiece material softens significantly. However, it is not melted and, therefore, not a liquid in the classical sense. From a modeling point of view, the problem can be addressed as both a fluid or a solid mechanics problem. Although it is technically possible to use a Lagrangian frame of reference (FoR) to solve the (Navier-)Stokes equations of fluid dynamics and to use an Eulerian FoR to solve Navier's equation of solid mechanics – in the context of FSW, the FoR dictates the overall approach. Using a Lagrangian FoR means treating FSW as a solid mechanics problem, and using an Eulerian FoR means treating FSW as a fluid mechanics problem.

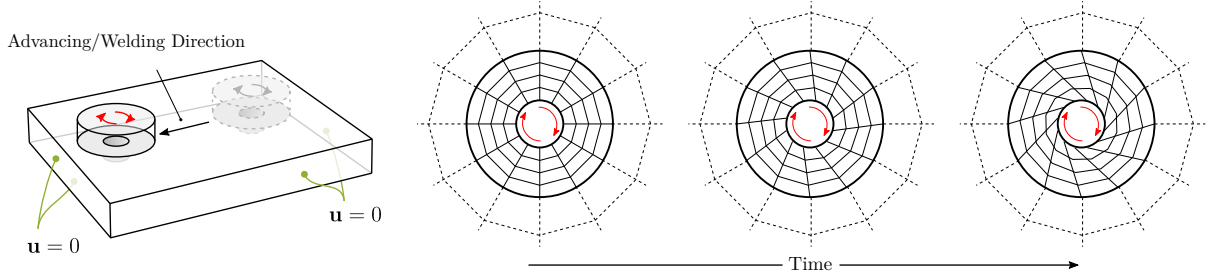
Lagrangian In a Lagrangian FoR, FSW is treated as a solid mechanics problem, where the primary unknowns are displacements and pressure, the latter being mandatory to allow for fully isochoric material deformation. While elastic deformations in the base material are non-isochoric, the continuous plasticization in the thermo-mechanically affected zone (TMAZ) and stir zone (SZ) is an isochoric process (see Sec. 2.1.1).

The constitutive model required for Lagrangian FSW simulations is thermo-elasto-visco-plastic, capturing both temperature-dependent elastic deformations and temperature-dependent isochoric visco-plastic deformations. With the mesh fixed to material particles, this approach allows precise tracking of material movement and enables residual stress analysis – an important aspect of FSW (see Par. *Joint Properties*).

However, contact modeling between the tool and workpiece is necessary, with friction laws based on contact pressure (e.g., Coulomb friction). Modeling contact can be a difficult task that can cause stability problems. However, a benefit of this approach is that the tool can be of arbitrary shape, and tilting the tool is no problem.

The advancement of the tool is modeled in a straightforward manner, and homogeneous Dirichlet boundary conditions are applied as shown in Fig. 4.11 (left). The Lagrangian approach is technically capable of simulating all three phases of FSW (plunging, welding, and extraction). However, the non-linear iterative process does not converge to a steady solution, even if the physical process has reached a steady state during the welding phase.

The major disadvantage of the Lagrangian framework is the severe mesh distortions that occur (Fig. 4.11, right). The reason is the strain localization close to the tool. Only a few degrees of tool rotation completely distort the interface elements. Frequent re-meshing is unavoidable and becomes computationally expensive. Even worse, frequent projection of unknowns and, especially, historical variables can harm the solution accuracy.



Isolated examples of the Lagrangian framework being used to simulate FSW are Guerdoux et al. [174], who concluded that the method may not be optimal for the translation phase, and Bussetta et al. [175], who compared both solids and fluids approaches and concluded that, while yielding comparable solutions, the fluids approach is computationally more efficient.

Eulerian In an Eulerian FoR, FSW is treated as a fluid mechanics problem, where the primary unknowns are velocity and pressure. Given that the plastic flow of metals drives the process, the incompressibility constraint must also be enforced.

By neglecting the elastic response entirely, the fluid approach simplifies the constitutive model, replacing the thermo-elasto-visco-plastic model with a thermo-visco-plastic one. Common visco-plastic models include the Sheppard-Wright model [176], the (regularized) Bingham model [177], the Carreau model [178], and the Norton-Hoff model [179, 180]. These models can describe a non-linear stress-strain rate relationship, with parameters dependent on temperature. A key advantage of thermo-visco-plastic models is the relatively easy material characterization. The required parameters can be obtained through simple torsion tests performed at various temperatures and strain rates.

Another advantage of the Eulerian approach is the absence of contact modeling. Interaction between the tool (*i.e.*, a moving wall) and the workpiece material (*i.e.*, the fluid) is modeled as viscous friction, independent of normal contact pressure. In 2D, the problem resembles a *flow-over-rotating-cylinder* scenario, and in 3D, it is a very similar problem except for a more complex shape of the “cylinder”.

Unlike the Lagrangian approach, the mesh remains undeformed, and the tool remains stationary. The advancement is modeled by material inflow on one side (inlet) and outflow on the opposite side (outlet), using homogeneous Dirichlet boundary conditions on the remaining faces (Fig. 4.12, left).

However, the Eulerian approach has two main limitations. First, the entire tool (shoulder and pin) must have axial symmetry[†] to keep the mesh unchanged. Second, like all fluid-based frameworks, it cannot predict residual stresses due to neglecting elastic strains.

A major advantage of the Eulerian framework is that the simulations converge to steady-state solutions, making them particularly useful for predicting temperature and pressure fields, as well as reaction forces

Figure 4.11: Lagrangian framework

[174]: Guerdoux et al. (2004), “Numerical simulation of the friction stir welding process using both Lagrangian and Arbitrary Lagrangian Eulerian formulations”

[175]: Bussetta et al. (2013), “Comparison of a Fluid and a Solid Approach for the Numerical Simulation of Friction Stir Welding with a Non-Cylindrical Pin”

[176]: Sheppard et al. (1979), “Determination of flow stress: Part 1 constitutive equation for aluminium alloys at elevated temperatures”

[177]: Santos et al. (2011), “Numerical approximations for flow of viscoplastic fluids in a lid-driven cavity”

[178]: Carreau (1972), “Rheological Equations from Molecular Network Theories”

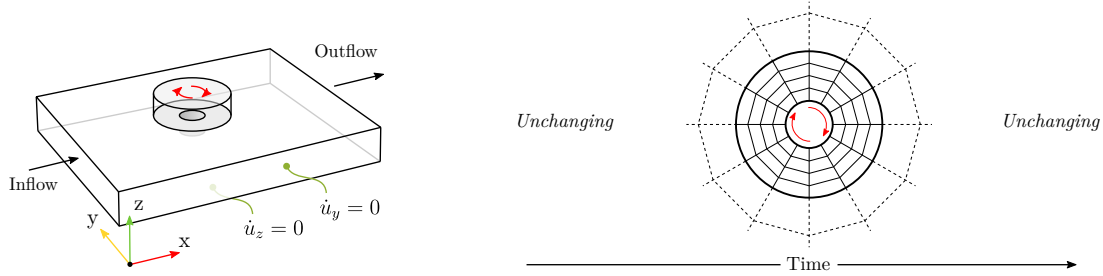
[179]: Norton (1929), *The Creep of Steel at High Temperatures*

[180]: Hoff (1954), “Approximate analysis of structures in the presence of moderately large creep deformations”

[†] also called *rotational symmetry of infinite order*

and torque during welding. However, they cannot simulate plunging or extraction phases. Tilting can be modeled without any problem.

While the velocity field accurately describes material flow, visualizing it intuitively can be challenging. However, streamlines and particle tracing methods can be employed to present the material flow in a more “Lagrangian” manner.



Examples of an Eulerian framework being used include works by De Vuyst and D’Alvise [181], who used their model to predict heat generation and material flow around simple tool geometries, Assidi and Fourment [182], who studied different friction laws and parameters, and Saracibar et al. [33] and Agelet de Saracibar et al. [183], who pioneered the use of VMS stabilized methods to simulate FSW problems.

Arbitrary-Lagrangian-Eulerian (ALE) The main drawback of the Eulerian framework is its limitation to axisymmetric tools. As discussed in Par. *Tool Shape* and shown in Figs. 4.6 and 4.7 (shoulder and pin types), a wide variety of tools is used in the industry and the overwhelming majority is threaded and/or has flutes or flats, respectively. Furthermore, the majority of shoulder features do not satisfy axial symmetry either.

The solution for this problem is the Arbitrary-Lagrangian-Eulerian (ALE) framework. This framework is still based on the fluids approach. However, unlike the Eulerian framework, it allows for mesh movement. As the tool rotates, the mesh is consequently deformed, as shown in Fig. 4.13. The deformed mesh is obtained by solving an elasticity or Poisson problem to which the tool position acts as a Dirichlet boundary condition.

On the one hand, the ALE framework overcomes a main limitation of the Eulerian framework; however, on the other hand, the old problem of the Lagrangian approach re-appears: strongly deformed meshes. As a result, re-meshing is required, and solutions must be mapped between meshes. Although it seems like nothing much has been gained, there are two noteworthy differences between the ALE and the Lagrangian frameworks with respect to re-meshing. First, since the mesh distortions are not fixed to the material movement, the distortions are less localized at the tool-workpiece interface. Therefore, re-meshing is not as frequent. Second, since thermo-visco-plastic material laws do not rely on history variables, the mapping between meshes is less prone to deteriorate the solution.

Figure 4.12: Eulerian framework

[181]: De Vuyst et al. (2007), “Material Flow Around a Friction Stir Welding Tool: Development of a Thermo-Fluid Code”

[182]: Assidi et al. (2009), “Accurate 3D friction stir welding simulation tool based on friction model calibration”

[33]: Saracibar et al. (2014), “Computational Modeling and Sub-Grid Scale Stabilization of Incompressibility and Convection in the Numerical Simulation of Friction Stir Welding Processes”

[183]: Agelet de Saracibar et al. (2010), “A Computational Model for the Numerical Simulation of FSW Processes”

Apart from the non-axisymmetric tools, the same possibilities and restrictions apply for the ALE framework as for the Eulerian framework. This means the constitutive law is thermo-visco-plastic, material flow visualization is realized through stream lines and particle tracing, explicit modeling of tilting is possible and residual stresses cannot be predicted.

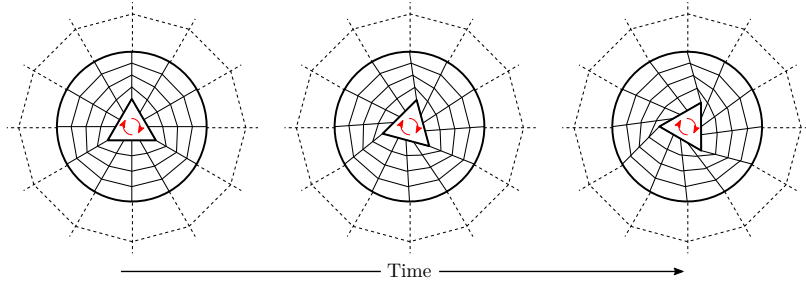


Figure 4.13: ALE framework

The ALE approach is used by most researchers nowadays and examples are studies conducted by Guerdoux and Fourment [184], who simulated all three phases of the FSW process. They treated the top surface as a free surface, thus fully modeling the flash and permitted tool and workpiece to dynamically loose and recover contact. Further examples are works by Myung et al. [185] who investigated the coupling of simulation-based temperature histories into a strength prediction model, which is formulated on the basis of precipitation kinetics and precipitate-dislocation interaction.

Apropos Kinematic Framework (AKF) The previous frameworks are all general concepts that can be used for a number of different applications. However, the following approach was developed by Chiumenti et al. [32] and Dialami et al. [186] specifically for FSW. The idea is to divide the computational domain into three parts: (1) the rigid tool, modeled in a Lagrangian framework; (2) a circular zone surrounding the tool, roughly the size of the HAZ, modeled in an ALE framework; and (3) the rest of the workpiece, modeled in an Eulerian framework.

Both the tool and HAZ rotate in unison at a constant speed (the rotational velocity) around the tool's axis of rotation. It is crucial to understand that the elements remain undeformed. The tool's elements do not deform since they are assumed to be rigid, and the HAZ's elements remain undeformed because the mesh movement (independent from the material flow!) is a purely circular, rigid motion. The tool and HAZ subdomains are effectively decoupled to allow for jumps in material velocity, pressure, and temperature. However, a heat flux is modeled between the two subdomains to allow for thermal contact via conduction.

The interface between the HAZ (ALE) and the rest of the workpiece (Eulerian) is of a different nature. Here, all fields (*material* velocity, pressure, temperature) have to be continuous, while the *mesh* velocities are different: the Eulerian part does not move while the ALE part rotates. To solve this problem, all interface nodes are duplicated, a *node-to-node link* is established, and a coupling between the two subdomains – similar to that of a periodic boundary condition – is achieved. The setup is shown in Fig. 4.14.

[184]: Guerdoux et al. (2009), “A 3D numerical simulation of different phases of friction stir welding”

[185]: Myung et al. (2020), “Probing the Mechanism of Friction Stir Welding with ALE Based Finite Element Simulations and Its Application to Strength Prediction of Welded Aluminum”

[32]: Chiumenti et al. (2013), “Numerical modeling of friction stir welding processes”

[186]: Dialami et al. (2013), “An apropos kinematic framework for the numerical modeling of friction stir welding”

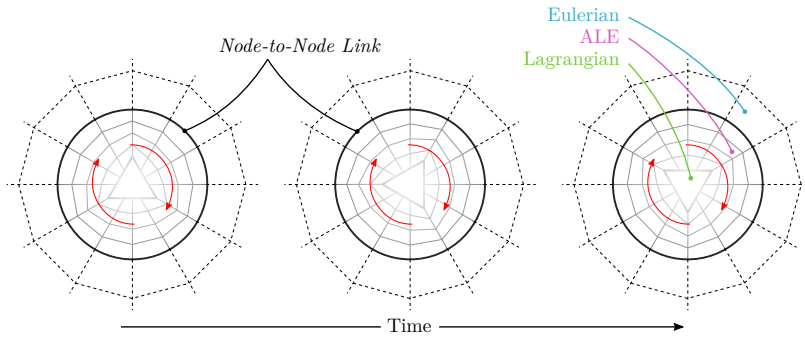


Figure 4.14: AKF framework

The AKF allows for non-axisymmetric tools, thus maintaining the benefits of the ALE framework while also preserving the undeformed elements from the Eulerian framework.

However, there is one significant limitation to the AKF. Since the HAZ (ALE) subdomain has to rotate in unison with the (Lagrangian) tool, no tilt angles are permitted. As shown in Fig. 4.15, it is impossible to match the interface between the Eulerian and ALE subdomain if the normal vector of the workpiece $\mathbf{n}$ is not aligned with the tools rotation axis $\mathbf{k}$, *i.e.*, if the tilt angle is larger than 0° . To address this issue, Dialami, Cervera, and Chiumenti [187] modified the friction law to account for the effect introduced by the tilt angle. However, it remains desirable to account for the tilt angle using geometric modeling rather than by modifying the friction law.

[187]: Dialami et al. (2018), “Effect of the Tool Tilt Angle on the Heat Generation and the Material Flow in Friction Stir Welding”

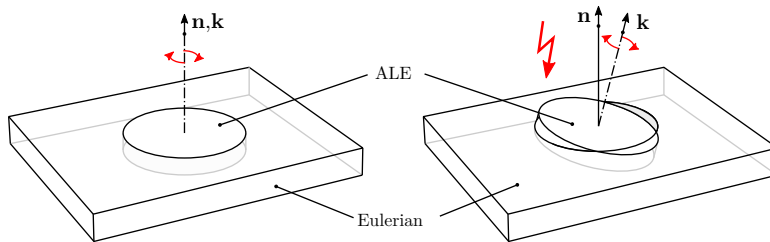


Figure 4.15: AKF framework
[188]: Dialami et al. (2017), “Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation”

[189]: Dialami et al. (2017), “A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding”

[190]: Dialami et al. (2017), “Experimental Validation of an FSW Model with an Enhanced Friction Law: Application to a Threaded Cylindrical Pin Tool”

[191]: Feulvarch et al. (2013), “A simple and robust moving mesh technique for the finite element simulation of Friction Stir Welding”

[192]: Robe et al. (2019), “R-ALE simulation of heat transfer during friction stir welding of an AA2xxx/AA7xxx joint on a large process window”

[193]: Bussetta et al. (2016), “Two 3D thermomechanical numerical models of friction stir welding processes with a trigonal pin”

Apart from the tilting restriction and the possibility of simulating non-axisymmetric tools, the AKF framework has the same possibilities and restrictions as the ALE and Eulerian frameworks. This means the constitutive law is thermo-visco-plastic, material flow visualization is realized through streamlines and particle tracing, and residual stresses cannot be predicted. The framework was deliberately developed to simulate the welding phase. Plunging and extraction cannot be simulated.

Examples of the AKF being used are works by Dialami et al. [188, 189] and Dialami et al. [190] who – within the AKF framework – developed enhanced friction models to better predict transversal forces and method to reach a steady state solution with less computational effort. Other examples include works by Feulvarch, Roux, and Bergheau [191] and Robe et al. [192], who used their model to predict the temperature evolution of dissimilar aluminum joints. Their method is called Rigid-ALE, though the fundamental approach is the same. Bussetta et al. [193] follow a similar approach. However, they introduce a transition zone instead of directly coupling the ALE and Eulerian domains.

Non-body fitted methods The most important phase of FSW is the actual welding phase. For this reason, and because of their enhanced computational efficiency, fluids-based frameworks are generally preferred. However, none of the above approaches can provide all of the following features simultaneously:

1. Undeformed mesh/elements,
2. Geometrical modeling of tilting,
3. Non-axisymmetric tools.

The absence of a fluids-based framework that simultaneously addresses all the above points motivates using non-body fitted methods. As introduced in Sec. 3 (*Unfitted FEM / Embedded Domain Methods*), these methods enable simulations on a structured background mesh, eliminating the need for the computational mesh to conform to the complex and moving boundary of the embedded geometry.

In this approach, the workpiece is represented as a Cartesian box and meshed in a fully structured manner. The tool, defined by a CAD file (e.g., STL, STEP, or IGES), is embedded into the domain at a user-specified position and orientation – allowing for tilting – while tool tractions and frictional heating are applied on the newly created immersed boundary. If the tool lacks axial symmetry, its angular position is updated at each time step according to the rotation speed, as illustrated in Fig. 4.16. Crucially, this is still possible for non-zero tilt angles. The Stokes flow equations are formulated within a purely Eulerian framework, as no mesh deformation occurs.

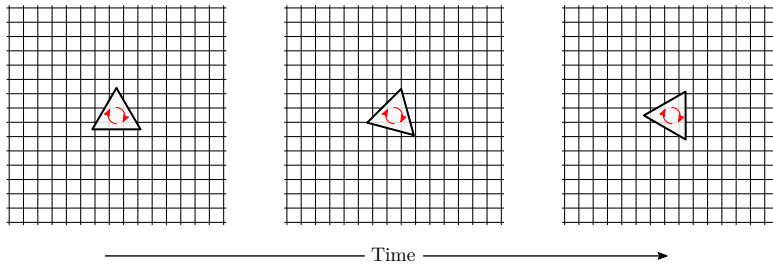


Figure 4.16: EDM framework

Ultimately, the non-body-fitted approach integrates all the advantages of other fluids-based frameworks. Its only limitations are those inherent to the fluids-based approach itself: the absence of native material flow visualization (which can be addressed by particle tracing) and the inability to predict residual stresses.

The application of non-body-fitted methods to FSW simulation is a novel concept. An isolated example is the work of Janoske, Bürger, and Geiger [194], which employed an immersed boundary method (IBM) within a fluids-based framework to predict the temperature evolution and compare the results to experimental data.

Fig. 4.17 presents the progression of numerical FSW analysis. Tab. 4.18 summarizes the various approaches, detailing their characteristics, advantages, and disadvantages.

[194]: Janoske et al. (2019), “Modelling of friction stir welding processes with a coupled immersed boundary-/volume of fluid-approach”

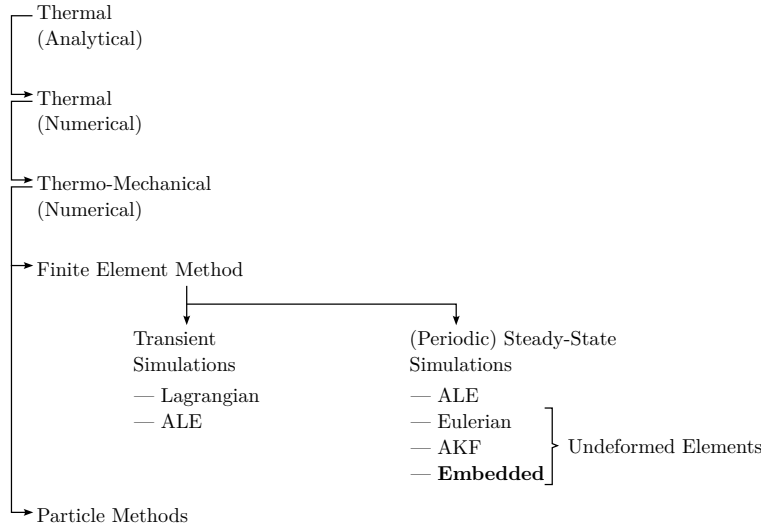


Figure 4.17: Evolution of analytical and numerical methods to analyze FSW processes

	Lagrangian	Eulerian	ALE	AKF	EDM
General Approach	Solids	Fluids	Fluids	Fluids	Fluids
Primary Unknowns	$\mathbf{u}, p, T$	$\mathbf{u}, p, T$	$\mathbf{u}, p, T$	$\mathbf{u}, p, T$	$\mathbf{u}, p, T$
Material Model	TEVP	TVP	TVP	TVP	TVP
Elastic Strains	✓ Yes	✗ No	✗ No	✗ No	✗ No
Residual Stresses	✓ Yes	✗ No	✗ No	✗ No	✗ No
Friction Modeling	Coulomb	Newton	Newton	Newton	Newton
Advancement Modeling	Tool Movement	Inflow/Outflow	Both possible	Inflow/Outflow	Both possible
Tool Geometry	Arbitrary	Rotational Sym.	Arbitrary	Arbitrary	Arbitrary
Undistorted Elements	✗ No	✓ Yes	✗ No	✓ Yes	✓ Yes
Material Flow Viz.	Native	Particle Tracing	Particle Tracing	Particle Tracing	Particle Tracing
Tilt Modeling	✓ Yes	✓ Yes	✓ Yes	✗ No	✓ Yes
Plunging Phase	✓ Yes	✗ No	✗ No	✗ No	✓ Yes
Welding Phase	✓ Yes	✓ Yes	✓ Yes	✓ Yes	✓ Yes
Extraction Phase	✓ Yes	✗ No	✗ No	✗ No	✗ No

4.3 Finite Element Framework

The last section motivated the use of non-body-fitted methods for simulating FSW. This is a very novel approach, and to the best of current knowledge, the only other work that similarly addresses FSW is the work by Janoske, Bürger, and Geiger [194].

However, their approach did not distinguish between the workpiece (*i.e.*, outside the immersed boundary) and the pin-tool (*i.e.*, inside the immersed boundary). Instead, only the interface between the workpiece and the tool was treated as a non-mesh-conforming boundary and integrated accordingly. As discussed in Sec. 3.2.4 (*Small Cut Stabilization*), this method avoids instability issues but sacrifices accuracy, as both sides of the immersed boundary remain coupled and, consequently, influence one another. Furthermore, works by Dialami, Cervera, and Chiumenti [187], Dialami et al. [188], and Dialami et al. [190] have shown that the assumption of a *full-stick* condition leads to an underestimation of lateral forces and that more advanced friction laws are necessary.

This work develops a comprehensive, high-performance computing frame-

Figure 4.18: Overview FSW Frameworks

TEVP: thermo-elasto-visco-plastic;
TVP: thermo-visco-plastic

[194]: Janoske et al. (2019), “Modelling of friction stir welding processes with a coupled immersed boundary-/volume of fluid-approach”

[187]: Dialami et al. (2018), “Effect of the Tool Tilt Angle on the Heat Generation and the Material Flow in Friction Stir Welding”

[188]: Dialami et al. (2017), “Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation”

[190]: Dialami et al. (2017), “Experimental Validation of an FSW Model with an Enhanced Friction Law: Application to a Threaded Cylindrical Pin Tool”

work for simulating FSW within a fluids-based approach. Employing CutFEM effectively and accurately handles complex and moving tool boundaries while leveraging previously developed, enhanced friction laws to achieve accurate predictions of temperature evolution, reaction forces, and torque.

The following section discusses the selected *Frame of Reference* (Sec. 4.3.1), *Boundary Conditions* (Sec. 4.3.3), and *Finite Element Approximation* (Sec. 4.3.4) on a general level. For a more detailed account, see Venghaus et al. [195][Appx. C]. The sections *Concave Tools* (Sec. 4.3.5) and *Surface Normals* (Sec. 4.3.6) cover further development which was not included in the publication.

[195]: Venghaus et al. (2025), “Embedded technology for enhanced modeling of Friction Stir Welding processes”

4.3.1 Frame of Reference

As indicated in Tab. 4.18, the non-body fitted, fluids-based framework can account for the advancement of the tool in two ways, as shown in Fig. 4.19. One possibility is to prescribe homogeneous Dirichlet boundary conditions on all faces and move the tool along the weld line (left). The tool movement can be easily facilitated by updating the position of the CAD file (both translatory and angular position) each time step. The other option is to keep the tool in place, only update the angular position, and account for the advancing of the tool by prescribing an inflow velocity at the inlet of the domain (right).

Since the entire simulation is carried out in a fluids-based framework, the goal is to reach a steady-state (or quasi-steady-state) solution. For this reason, the second option (material inflow modeling) is preferred.

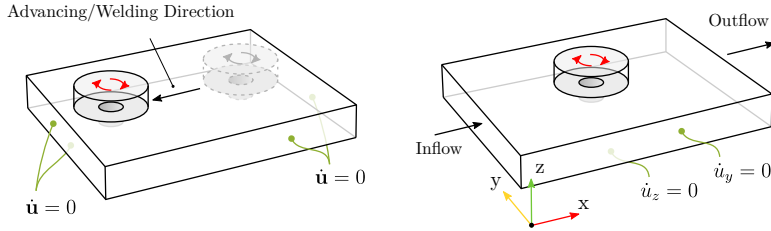
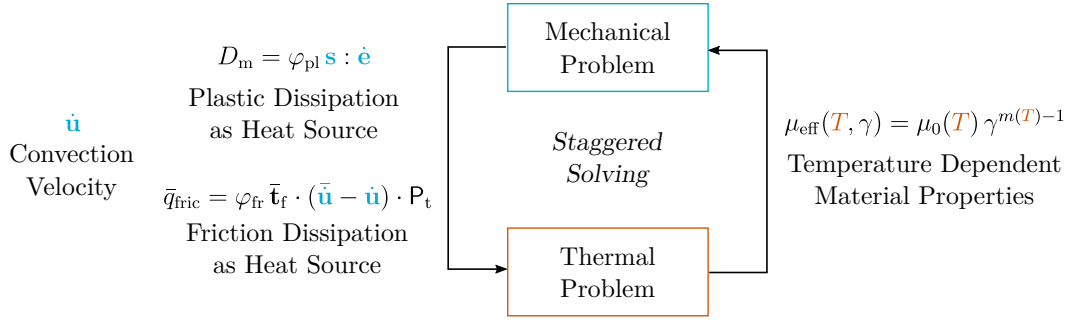


Figure 4.19: Tool advancement (left) vs. material inflow modeling (right)

4.3.2 Thermo-Mechanical Coupling

FSW is typically modeled as a coupled thermo- (*convection-diffusion*) mechanical (*Stokes flow*) problem. The coupling is bi-directional, as shown in Fig. 4.20. Plastic dissipation, friction dissipation, and material velocity are the results of the mechanical problem and influence the thermal problem as source terms and convective velocity, respectively. The resulting temperature field influences the mechanical problem by altering the mechanical properties, such as viscosity and the severity of the shear-thinning effect.



Since the interaction between the mechanical and the thermal problem is weak, it is sufficient to solve the coupled system in a staggered manner. Furthermore, since the goal is to reach a steady state (for axisymmetric tools) or periodic steady state solution (for non-axisymmetric tools), it is sufficient to solve each part of the coupled problem only once per time step.

Figure 4.20: Thermo-mechanical coupling

The strong form of the Stokes flow problem reads

$$\nabla \cdot \mathbf{s} + \nabla p + \mathbf{b} = \mathbf{0} \quad (4.1)$$

$$\nabla \cdot \dot{\mathbf{u}} = 0, \quad (4.2)$$

and the strong form of the convection-diffusion equation reads

$$\rho_0 c \left(\frac{\partial T}{\partial t} + \dot{\mathbf{u}} \cdot \nabla T \right) + \nabla \cdot \mathbf{q} = D_m \quad \text{with} \quad \mathbf{q} = -k \nabla T. \quad (4.3)$$

The unknowns are $\dot{\mathbf{u}}$, the velocity; p , the pressure; and T , the temperature. The remaining variables are $\mathbf{s}$, the deviatoric stress; $\mathbf{b}$, the body forces; ρ_0 , the density; c , the specific heat capacity; k , the thermal conductivity; $\mathbf{q}$, the heat fluxes; and $D_m = \varphi_{pl} \mathbf{s} : \dot{\mathbf{e}}$, the source term, i.e., the fraction of the plastic dissipation that is converted into heat. The plastic dissipation is defined as the double contraction of strain rate and stress (i.e., $s_{ij} \dot{e}_{ij}$), and the proportionality factor φ_{pl} determines how much of the plastic dissipation is accounted for as a heat source for the thermal problem.

The non-Newtonian material flow is described using a Norton-Hoff [179, 180] material law, which models the strain rate-stress relation as a power law.

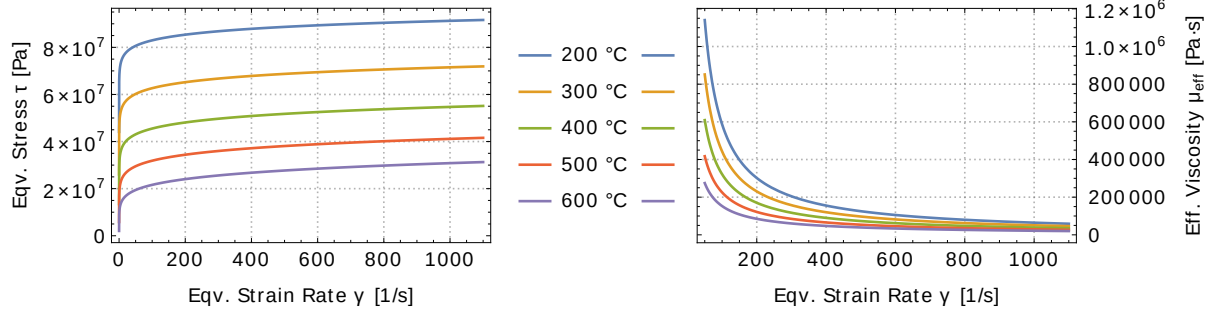
$$\mathbf{s} = 2\mu_{eff} \dot{\mathbf{e}} = 2\mu_{eff} \nabla^s \dot{\mathbf{u}} \quad \text{with} \quad \mu_{eff}(T, \gamma) = \mu_0(T) \gamma^{m(T)-1}, \quad (4.4)$$

where $\dot{\mathbf{e}} = \nabla^s \dot{\mathbf{u}}$ is the deviatoric strain rate, $\nabla^s \dot{\mathbf{u}}$ is the symmetric gradient of the velocity, $\gamma = \sqrt{2} \|\dot{\mathbf{e}}\|$ is the equivalent strain rate, $\tau = \mu_{eff} \gamma$ is the equivalent stress, and μ_0 is the consistency parameter. Finally, m is the so-called flow index. Shear-thinning behavior is observed for $m < 1$, and Newtonian behavior is recovered for $m = 1$.

The model's temperature dependency further influences its non-linear behavior. Higher temperatures lead to lower viscosities and a weaker shear thinning effect, as illustrated in Fig. 4.21.

[179]: Norton (1929), *The Creep of Steel at High Temperatures*

[180]: Hoff (1954), "Approximate analysis of structures in the presence of moderately large creep deformations"



Due to the highly non-linear nature of the presented material law, a linearized version of the Norton-Hoff law was proposed by Dialami et al. [189]. See also Venghaus et al. [64][Appx. A]. The power law is converted to a piecewise linear function by evaluating the original law at strain rates of 1.0, 10.0, 100.0, and 1000.0 1/s and linearly interpolating the results, as seen in Fig. 4.22.

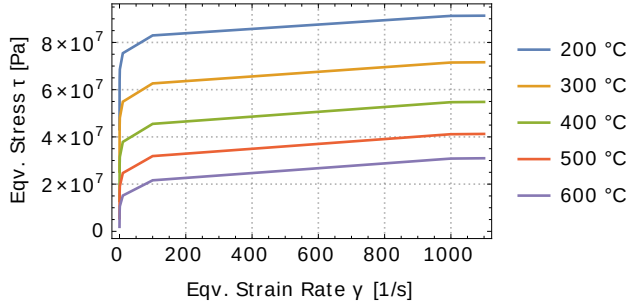


Figure 4.21: Temperature and strain rate dependent effective viscosity and equivalent stress

[189]: Dialami et al. (2017), “A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding”

[64]: Venghaus et al. (2023), “An accurate approach to simulate friction stir welding processes using adaptive formulation refinement”

Figure 4.22: Linearized material law

4.3.3 Boundary Conditions

The boundary conditions of the mechanical problem are defined as

$$\dot{\mathbf{u}} = \bar{\dot{\mathbf{u}}} \quad \text{on} \quad \Gamma_u \quad \text{and} \quad (4.5)$$

$$\mathbf{s} \cdot \mathbf{n} + p\mathbf{n} = \bar{\mathbf{t}} \quad \text{on} \quad \Gamma_t, \quad (4.6)$$

and as

$$T = \bar{T} \quad \text{on} \quad \Gamma_T, \quad \text{and} \quad (4.7)$$

$$\mathbf{q} \cdot \mathbf{n} = \bar{q} \quad \text{on} \quad \Gamma_q, \quad (4.8)$$

for the thermal problem.

Fig. 4.23 (left) shows the application of boundary conditions on the exterior boundaries, *i.e.*, the mechanical constraints $u_y = u_z = 0$, the advancing velocity $u_x = u_{\text{adv}}$, the heat transfer to the environment q_{lat} , q_{back} , and q_{top} , and the fixed inflow temperature $T = T_{\text{in}}$.

The interaction of the tool and workpiece can be split into two parts:

$$\bar{\mathbf{t}} = \bar{\mathbf{t}}_{\text{di}} + \bar{\mathbf{t}}_{\text{fr}}. \quad (4.9)$$

The first part, $\bar{\mathbf{t}}_{\text{di}}$, is associated with material being displaced by the rotating tool and occurs normal to the tool's surface ($\bar{\mathbf{t}}_{\text{di}} \parallel \mathbf{n}$). The

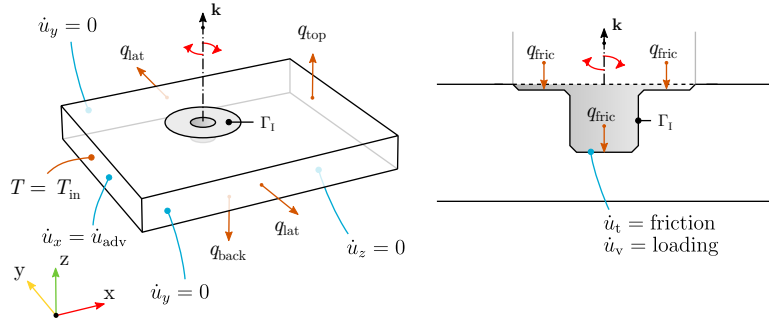


Figure 4.23: Thermal and mechanical boundary conditions on the exterior boundaries (left) and the tool-workpiece interface Γ_I , visualized as a cross section of the domain (right)

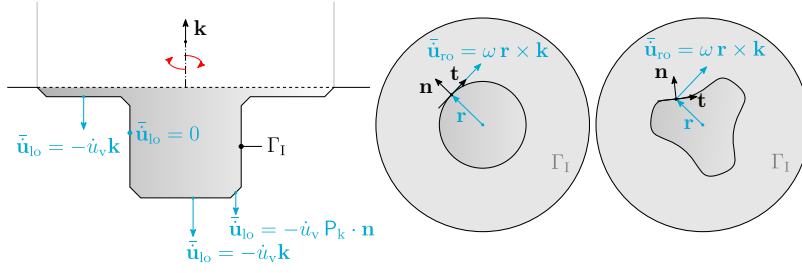


Figure 4.24: Kinematics and exterior normals on the tool-workpiece interface Γ_I : cross section (left), top view for axisymmetric pin (middle), and top view for non-axisymmetric pin (right)

second part, $\bar{\mathbf{t}}_{\text{fr}}$, results from friction, which occurs in-plane with the tool's surface ($\bar{\mathbf{t}}_{\text{fr}} \perp \mathbf{n}$).

For easier writing of the following traction definitions, three operators (normal, tangential, and axial) are introduced as

$$\mathbf{P}_n = \mathbf{n} \otimes \mathbf{n}, \quad (4.10)$$

$$\mathbf{P}_t = \mathbf{1} - \mathbf{P}_n, \quad (4.11)$$

$$\mathbf{P}_k = \mathbf{k} \otimes \mathbf{k}, \quad (4.12)$$

where $\mathbf{k}$ denotes the tool rotation axis (see Fig. 4.24, left).

The displacement and friction tractions are defined as

$$\bar{\mathbf{t}}_{\text{di}} = b \, (\bar{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{P}_n, \quad (4.13)$$

$$\bar{\mathbf{t}}_{\text{fr}} = a^* \, (\bar{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{P}_t, \quad (4.14)$$

where a^* and b are penalty parameters, and $\bar{\mathbf{u}}$ is the prescribed velocity. As indicated in Fig. 4.24, the prescribed velocity $\bar{\mathbf{u}}$ consists of two parts: rotational movement of the tool and the vertical loading:

$$\bar{\mathbf{u}} = \underbrace{\omega \mathbf{r} \times \mathbf{k}}_{\bar{\mathbf{u}}_{\text{ro}}} - \underbrace{\dot{u}_v \mathbf{P}_k \cdot \mathbf{n}}_{\bar{\mathbf{u}}_{\text{10}}}. \quad (4.15)$$

The rotational velocity $\bar{\mathbf{u}}_{\text{ro}}$ is a function of the angular speed of the tool ω and the distance $\mathbf{r}$ to the tool's rotation axis $\mathbf{k}$. The vertical loading $\bar{\mathbf{u}}_{\text{10}}$ is introduced to account for the downforce aligned with $\mathbf{k}$.

The penalty for the material displacement b is chosen large enough to ensure $(\bar{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{P}_n \approx 0$.

The friction is accounted for by

$$a^* = a ||(\bar{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{P}_t||^{q-1}, \quad (4.16)$$

where a and q are the friction and sensitivity parameters, respectively. See Dialami et al. [188] and Venghaus et al. [195][Appx. C] for more details.

Remark 4.3.1. *Note that material displacement is only present for non-axisymmetric tools. For axisymmetric tools, all surface normals are orthogonal to the local rotational tool speed $\mathbf{n} \perp \bar{\mathbf{u}}_{\text{ro}}$, resulting only in frictional contact and no displacement.*

The heat transfer to the environment ($\bar{q}_{\text{lat}}$, $\bar{q}_{\text{top}}$, $\bar{q}_{\text{back}}$) is modeled as a Robin boundary condition. See Eq. (4.17), where $h(T)$ is the heat transfer coefficient, T is the temperature at the domain boundary, and T_{env} is the environment temperature.

$$\bar{q} = h(T) (T - \bar{T}_{\text{env}}) \quad (4.17)$$

The heat fluxes on Γ_I are caused by friction dissipation and are computed as

$$\bar{q}_{\text{fric}} = \varphi_{\text{fr}} \bar{\mathbf{t}}_f \cdot (\bar{\mathbf{u}} - \dot{\mathbf{u}}) \cdot \mathbf{P}_t. \quad (4.18)$$

They are based on the tangential components of the tractions and velocity differences, as introduced previously. The factor φ_{fr} determines the proportion of the dissipated energy that is converted to heat. See Dialami et al. [196] for a more detailed explanation.

4.3.4 Finite Element Approximation

A detailed derivation of the Finite Element approximation of the FSW problem is not necessary at this point since the mixed velocity-pressure (u/p) problem for non-body fitted meshes has been derived in the previous Section (*Unfitted FEM / Embedded Domain Methods*, Sec. 3).

The mechanical problem in weak form reads

[188]: Dialami et al. (2017), “Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation”

[195]: Venghaus et al. (2025), “Embedded technology for enhanced modeling of Friction Stir Welding processes”

[196]: Dialami et al. (2016), “Challenges in Thermo-mechanical Analysis of Friction Stir Welding Processes”

$$\begin{aligned}
& \int_{\Omega} \mu_{\text{eff}} \nabla^s \delta \dot{\mathbf{u}}_h : \nabla^s \dot{\mathbf{u}}_h \, dV \\
& + \int_{\Omega} \nabla \cdot \delta \dot{\mathbf{u}}_h p_h \, dV + \int_{\Omega} \delta p_h \nabla \cdot \dot{\mathbf{u}}_h \, dV \\
& - \sum_K \tau_u \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV \\
& + \sum_K \gamma_{1K} \int_{K(\Omega^{1\text{st}2})} \nabla \delta \dot{\mathbf{u}}_h \cdot \mathbf{P}_h^\perp [\nabla \dot{\mathbf{u}}_h] \, dV \\
& + \sum_K \gamma_{2K} \int_{K(\Omega^{1\text{st}2})} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h - \mathbf{b}] \, dV = \int_{\Gamma_I} \delta \dot{\mathbf{u}}_h \cdot b (\bar{\mathbf{u}} - \dot{\mathbf{u}}_h) \cdot \mathbf{P}_n \, dA \\
& + \int_{\Gamma_I} \delta \dot{\mathbf{u}}_h \cdot a^* (\bar{\mathbf{u}} - \dot{\mathbf{u}}_h) \cdot \mathbf{P}_t \, dA \\
& \text{with } \dot{\mathbf{u}}_h = \bar{\mathbf{u}} \text{ on } \Gamma_u, \text{ according to Fig. 4.23 (left).}
\end{aligned} \tag{4.19}$$

The thermal problem reads

$$\begin{aligned}
& \int_{\Omega} \delta T_h \rho_0 c \left(\frac{1}{\alpha} \frac{\partial T_h}{\partial t} + \dot{\mathbf{u}}_h \cdot \nabla T_h \right) \, dV + \int_{\Omega} \nabla \delta T_h \cdot k \nabla T_h \, dV \\
& + \sum_K \varsigma_K \int_{\Omega_K} \dot{\mathbf{u}}_h \cdot \nabla \delta T_h \cdot \mathbf{P}_h^\perp [\dot{\mathbf{u}}_h \cdot \nabla T_h] \, dV \\
& + \sum_K \varepsilon_K \int_{K(\Omega^{1\text{st}2})} \nabla \delta T_h \cdot \mathbf{P}_h^\perp [\nabla T_h] \, dV = \int_{\Omega} \delta T_h D_m \, dV \\
& + \int_{\Gamma_q} \delta T_h \mathbf{q}_h \cdot \mathbf{n} \, dA \\
& + \int_{\Gamma_I} \delta T_h \varphi_{\text{fr}} \bar{\mathbf{t}}_f \cdot (\bar{\mathbf{u}} - \dot{\mathbf{u}}_h) \cdot \mathbf{P}_t \, dA \\
& \text{with } T_h = \bar{T}_{\text{in}} \text{ on } \Gamma_T, \text{ and } \mathbf{q}_h \cdot \mathbf{n} = \bar{q} = h (T_h - \bar{T}_{\text{env}}) \text{ on } \Gamma_q, \text{ according to Fig. 4.23 (left).}
\end{aligned} \tag{4.20}$$

The parameter α is introduced to artificially accelerate the transient phase to achieve the steady state faster. This method was first introduced by Dialami et al. [189].

Harmonic Extension A 2D slice of the computational domain is displayed in Fig. 4.25. It is divided into the workpieces (physical subdomain, gray) and the tool (white). The iso-zero surface of the level set function is indicated in red. Assuming a rigid tool and neglecting any heat flux between the tool and workpiece, the subdomain representing the pin tool remains inactive – both for the mechanical and the thermal problem. Consequently, no solution is required in the corresponding elements, and the inactive subdomain is entirely disregarded by assuming zero for all unknowns.

Issues arise when the level set function is updated at each time step, as is necessary for non-axisymmetric tools. As a result, some FE nodes

[189]: Dialami et al. (2017), “A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding”

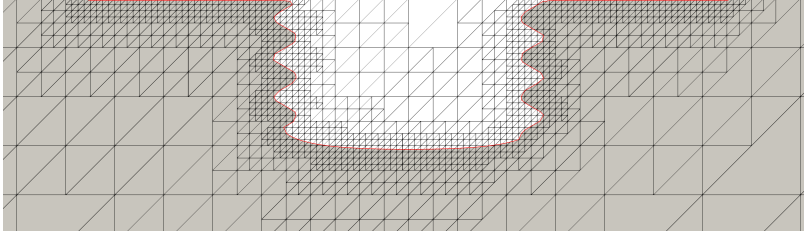


Figure 4.25: 2D slice of computational domain. Level set function highlighted in red.

transition from the inactive subdomain to the physical one. The absence of realistic initial conditions for these newly activated nodes negatively impacts convergence. To address this, the inactive subdomain is not completely ignored; instead, a harmonic extension is computed (see [197]). This is applied to both the thermal and mechanical problems, ensuring well-conditioned initial values for the non-linear iterative process.

The mechanical problem solved in the non-physical domain is a decoupled Poisson problem

$$\int_{\Omega} \mu \nabla \delta \mathbf{u}_h : \nabla \mathbf{u}_h \, dV + \int_{\Omega} \tau_u \nabla \delta p_h \cdot \nabla p_h \, dV = 0. \quad (4.21)$$

For the thermal problem the following Poisson problem is used to compute the harmonic extension

$$\int_{\Omega} k \nabla \delta T_h \cdot \nabla T_h \, dV = 0. \quad (4.22)$$

It is important to note that even when a harmonic extension is employed in the non-physical part, the physical subdomain needs to remain uncoupled from the non-physical one. This uncoupling is illustrated in Fig. 4.26. The treatment of the cut element was discussed in Sec. 3.2.2 (*Integration Method*). Now, the adjacent element on the non-physical side is of interest. Here, the problems shown in Eqs. (4.21) and (4.22) are solved. To prevent influencing the solution of the physical domain, all contributions to the *transition nodes* must be eliminated, as indicated in Fig. 4.26. However, the modification of the elemental stiffness matrix cannot happen in a symmetric manner as the physical side must still influence the non-physical side to be able to extend the physical solution harmonically.

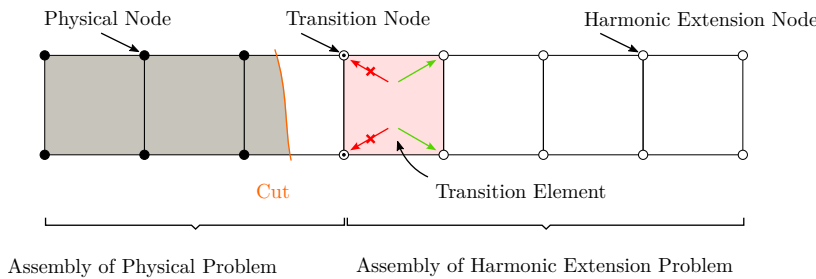


Figure 4.26: Harmonic extension and one-sided domain decoupling

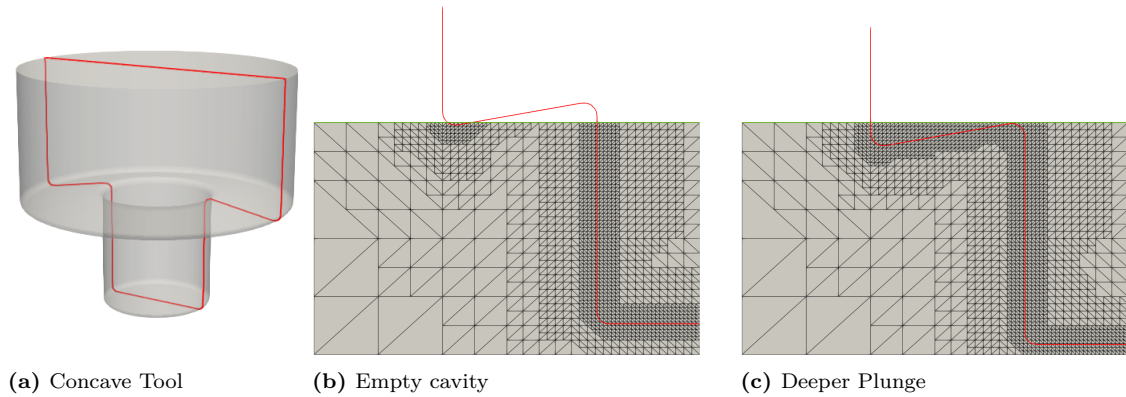
4.3.5 Concave Tools

As discussed in Sec. 4.1, Par. *Tool Shape*, concave shoulders (Fig. 4.27a) are commonly used, since they better trap the stirred material and reduce flashing. However, from a FEM modeling point of view, concave shoulders introduce a new challenge.

Since the workpiece is modeled as a rectangular box, the top surface is a simple plane in x - y . For explanatory purposes, it is assumed that the top surface is positioned at $z = 0$. If a tool is concave, the highest point of the shoulder is close to the pin, and the lowest point is at the outer brim of the shoulder. As a result, a cavity forms close to the pin if the tool is only plunged until the outer brim of the shoulder touches the workpiece.

In the real processes this cavity is filled with plasticized material and contact is established on the entire tool surface. However, in a Eulerian, fluids-based FE simulation, where mesh movement is not permitted, the cavity remains empty. Without further implementation, this situation can be addressed in two ways.

1. A realistic plunge depth is chosen, which leads to a slight penetration of the shoulder brim. As a result, the cavity close to the pin is left unfilled, and the contact area between the tool and the workpiece is significantly smaller. See Fig. 4.27b.
2. A unrealistically large plunge depth is chosen such that the highest point of the concave shoulder is below $z = 0$, resulting in a completely filled cavity. See Fig. 4.27c.



However, both options are unrealistic scenarios. A better way to address the issue is by adding additional elements above $z = 0$, filling the cavity, ensuring tool-workpiece contact on the entire shoulder surface, and simultaneously allowing for a realistic plunging depth.

The approach requires careful handling of the different physical and non-physical subdomains. All elements *inside* the iso-zero surface (depicted in red (//) in Fig. 4.29) belong to the tool and are subject to the harmonic extension problem, as described in Sec. 4.3.4, Par. *Harmonic Extension*. All elements depicted in blue (■) are also classified as non-physical. However, no harmonic extension is computed. The elements are simply disregarded and not assembled to the system; similarly to the deactivated DOF of the adaptive formulation refinement, they are

Figure 4.28: Different approaches to handling concave tools

not assembled in Sec. 2.5.5. The remaining elements belong to the physical subdomain.

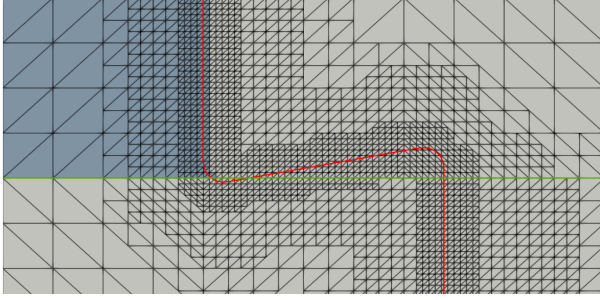


Figure 4.29: Excluded elements

Remark 4.3.2. *To ensure that the deactivated elements (■) do not affect the solution within the physical subdomain, they must be excluded during the assembly of the system. However, as discussed regarding the OSS stabilization in Sec. 2.4, Par. [Projections](#), and concerning the Ghost Penalty stabilization in Sec. 3.2.4, Par. [Ghost Penalty](#), projection and smoothing operations introduce a weak coupling between elements that do not share nodes but fall within the influence range of the smoothing.*

Therefore, it is also essential to exclude deactivated elements from projection and smoothing operations. Note that this also includes the computation of lumped mass matrices.

4.3.6 Surface Normals

The correct definition of boundary conditions on the level-set-defined boundary is crucial. Since the boundary conditions defined in Eqs. (4.13), (4.14) and (4.18) depend on normal and tangential projectors, $P_n = \mathbf{n} \otimes \mathbf{n}$ and $P_t = \mathbf{1} - P_n$, the accurate computation of the surface normal $\mathbf{n}$ is essential.

Sec. 3.2.2 covered the determination of integration points and weights. As a reminder, the discrete level-set function was computed as the (signed) shortest orthogonal distance of FE nodes to an embedded CAD geometry. Boundary-conform integration was then used to locate surface integration points and determine their weights. This method was used because of its high accuracy for linear elements and relatively low number of resulting integration points.

When a discrete level-set function intersects a linear tetrahedron, the resulting iso-zero surface forms either a triangle or a planar quadrilateral. In such cases, all integration points share a common normal vector computed from the nodal coordinates and the discrete level-set values. However, while true to the discrete level-set function, this normal does not necessarily match the corresponding normal from the original CAD geometry.

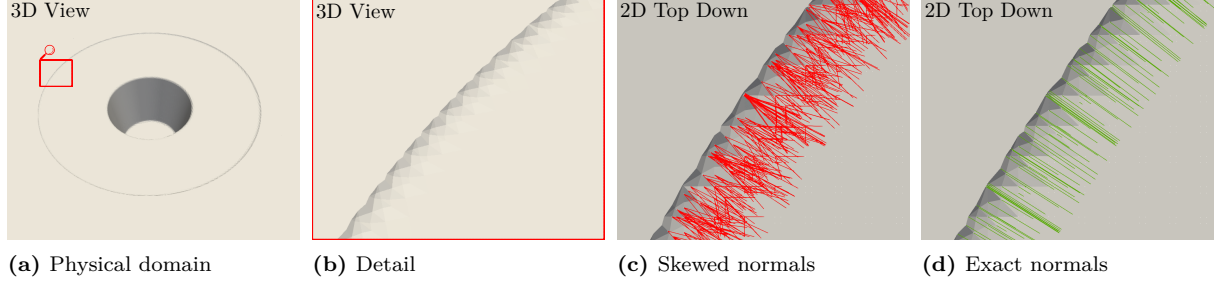


Fig. 4.31 illustrates the shoulder brim section of the physical domain. Due to the nature of linear elements and the structured mesh, the surface appears jagged, affecting the computed normals, as shown in Fig. 4.30c. For an axisymmetric tool, all interface normals should point towards the rotation axis and have no tangential component, *i.e.*, $\mathbf{n} \cdot \bar{\mathbf{u}} = 0$.

Figure 4.31: Recovery of exact normals

To correct this, the original normals are recovered from the CAD geometry through the following steps:

1. Compute the global coordinates of each surface integration point.
2. Orthogonally project each point onto the CAD surface.
3. Recover the CAD surface normal at the projected point.

Fig. 4.30d shows the corrected normals, pointing radially towards the rotation axis without tangential components.

4.4 Numerical Results

Forces evolution The reaction forces F_x , F_y , F_z and torque M are obtained according to Fig. 4.32. For each direction x (advancing), y (transversal), and z (vertical), as well as the torque with respect to the tool rotation axis $\mathbf{k}$, the value is recorded and saved. The same is done for the maximum temperature $T_{\max}$ in the physical domain. Fig. 4.34 shows their evolution over time. A total number of four rotations were simulated. Fig. 4.33 shows the respective tool shapes.

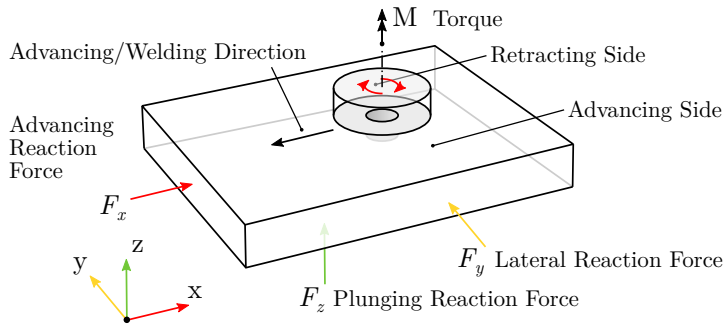


Figure 4.32: Definition of torque and advancing, lateral and plunging reaction forces

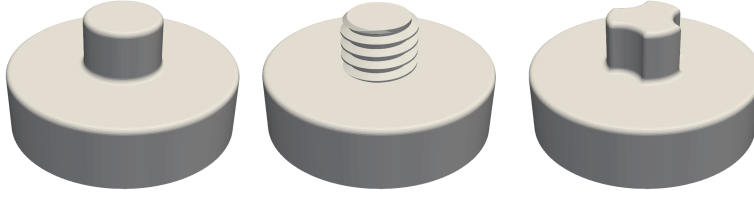


Figure 4.33: A set of different tools: cylindrical (left), threaded (middle) and trivex (right)

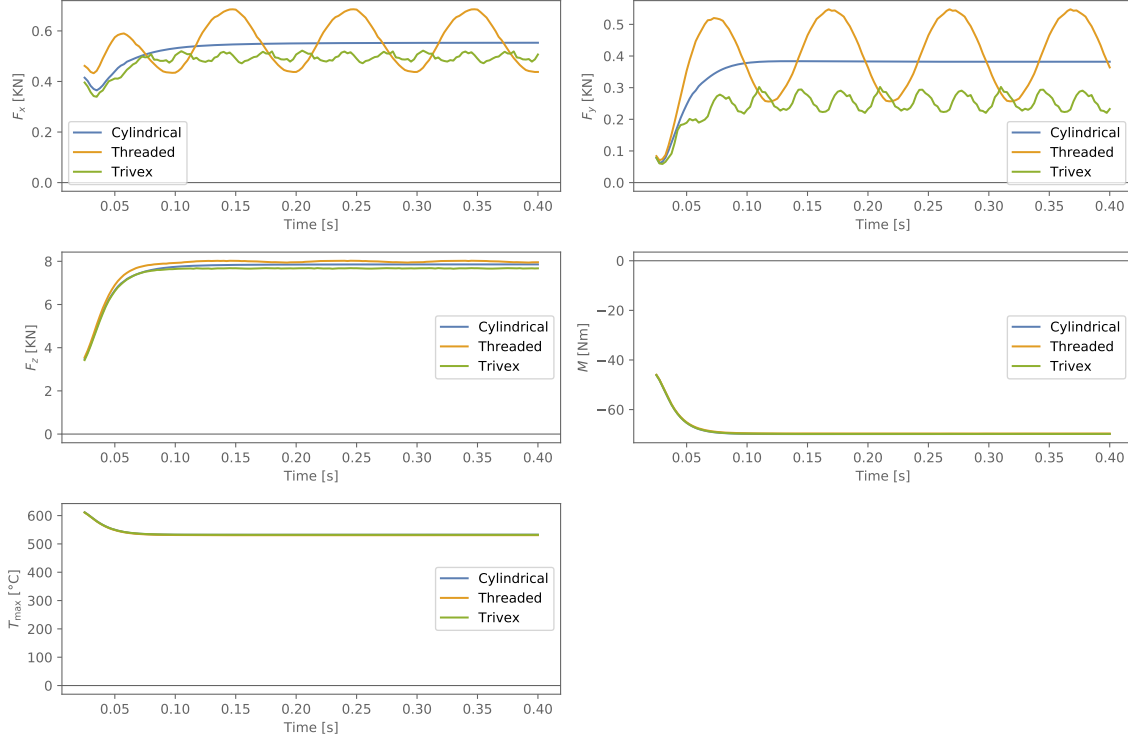
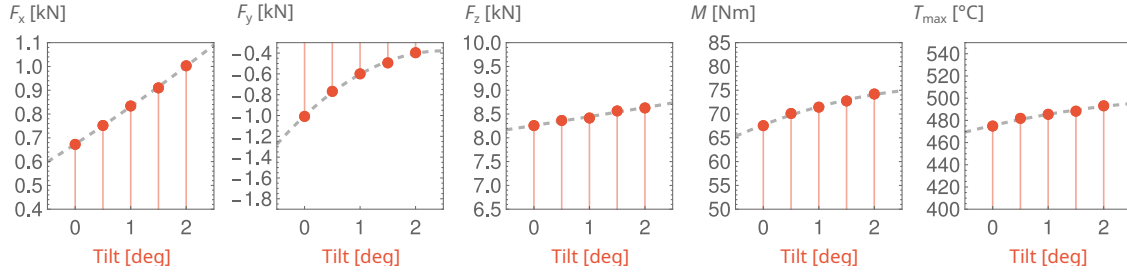


Figure 4.34: Forces, torque and maximum Temperature evolution over time

From the evolution of the forces over time, it becomes evident that, for the cylindrical tool, the forces converge to a stationary value. In contrast, for the non-axisymmetric tools, the forces exhibit oscillatory behavior. For the threaded tool, which has one-fold rotational symmetry (*i.e.*, a full 360° rotation is required to reproduce the same geometry), the oscillation frequency matches the tool's rotational frequency – one oscillation per rotation. For the trivex tool, which exhibits three-fold rotational symmetry (*i.e.*, a 120° rotation yields an identical geometry), the oscillation frequency is three times the rotational frequency. Typically, the vertical force F_z , torque M , and maximum temperature $T_{\max}$ do not oscillate regardless of the tool geometry.

Tilting Angle The driving motivation to develop an alternative to the Apropos Kinematic Framework was its limitation regarding non-zero tilt angles (see Fig. 4.15). Thanks to switching to a non-body fitted approach, the tilt angle is accounted for by simply adjusting the CAD's position before computing the discrete level-set function. Similar to the parameter study presented in Venghaus et al. [195][Appx. C], Fig. 4.35 shows the influence of the tilt angle on the reaction forces and maximum temperature.

[195]: Venghaus et al. (2025), “Embedded technology for enhanced modeling of Friction Stir Welding processes”



The influence of the tilt angle can be seen on all reaction forces, including torque and maximum temperature. Its impact is strongest on the advancing and transversal forces F_x and F_y . With an increasing tilt angle, the advancing forces rise. This effect can be attributed to the larger cross-section and an increased “angle of attack”. For the same reason, the vertical reaction forces F_z increase slightly with an increasing tilt angle.

The most significant part of the transversal reaction force is generated through slipping in the rear section of the tool. Under the leading edge, where the material is sticking to the tool, large transversal forces are generated. In the rear section, where the tool is slipping, smaller forces are generated. Since both forces act in opposite directions, the result is a transversal reaction force. The more slipping occurs in the rear part of the tool, the bigger the difference in transversal forces and the bigger the transversal reaction force. See Dialami et al. [188] and Venghaus et al. [195][Appx. C] for an in-depth explanation of these effects.

If a tilt angle is introduced, the rear part of the shoulder is plunged deeper into the workpiece, increasing its contact area. A larger contact area of the rear part of the tool counters the effect created by the rear slipping, decreasing the difference between front and rear transversal forces. Thus larger tilt angles lead to smaller transversal reaction forces F_y , as seen in Fig. 4.35.

The increased contact area can also be attributed to the increase in torque and maximum temperature for larger tilt angles.

Figure 4.35: Impact of tilting angle ($0^\circ \rightarrow 2^\circ$) on reaction forces, torque and maximum temperature

[188]: Dialami et al. (2017), “Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation”

[195]: Venghaus et al. (2025), “Embedded technology for enhanced modeling of Friction Stir Welding processes”

4.5 Bridging Research and Industrial Application

The transition from research-driven numerical methods to industrial applications often faces significant hurdles. While researchers focus on developing novel simulation techniques, engineers need robust, efficient, and user-friendly tools for real-world applications. Although research codes are often open-source, they can be challenging to use and integrate into industrial workflows. Commercial FEM codes such as Abaqus, ANSYS, or MSC Nastran are typically easier to operate but often lack novel numerical methods.

This challenge is particularly significant for FSW due to high computational demands and complex material modeling. Various advanced numerical techniques were introduced, such as employing different mixed

formulations for adaptive refinement, orthogonal subscale stabilization, non-body-fitted methods that interface directly with CAD data, and Ghost-Stabilization. All of these methods have been introduced because they offer higher accuracy and/or computational efficiency. However, none of these approaches are natively available in commercial FEM software. Even when utilizing programmable user interfaces, these methods are very complex to implement without having complete control over the entire FE code.

The task at hand is to bridge research and industrial application. This section addresses key aspects of this transition, including pre- and postprocessing and the implementation of said advanced methods in an HPC framework.

Preprocessing One of the most important aspects of FEM software accessibility is preprocessing. A streamlined and user-friendly way to define the simulation parameters is crucial to familiarize engineers with new simulation software quickly. For FSW, there are approximately 20 parameters, as shown in Tab. 4.1.

Processing	Dimensions	Materials/Friction	Output
Advancing Speed	Plate Dimensions	Viscosity	Output Frequency
Rotation Speed	Tool Geometry (CAD)	Flow Index	Required Output Fields
Plunging Depth	Tool Dimensions	Specific Heat Capacity	– Temperature
Tilt Angle	Mesh Type and Density	Conductivity	– Material Tracing
Rotation Direction	Simulation Time	Friction Coefficients	– Strain Rates...

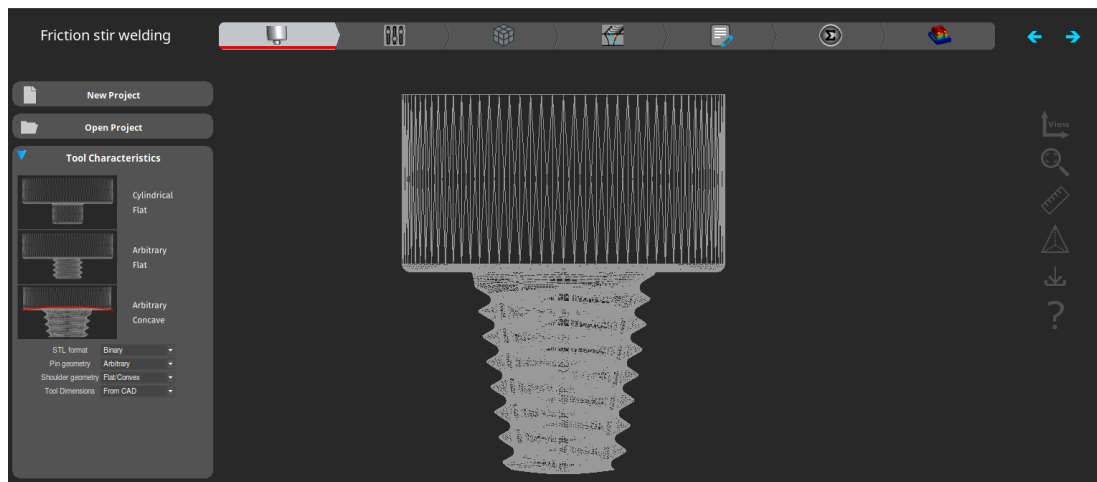
Typically, the geometry description and mesh generation are a time consuming part of preprocessing. Fortunately, using a non-body-fitted method, this aspect is already reduced to an absolute minimum. A rectangular box represents the workpiece, and the tool-workpiece interface is defined using a discrete level-set function which is directly computed from a CAD file, as described in Sec. 3.2.1, Par. *Implicit Geometry Description*.

To facilitate user-friendly preprocessing, a graphical user interface (GUI, Fig. 4.36) was developed to streamline the processes. After launching the GUI, the user can select a CAD file from the file system (STL/STEP/IGES), which is loaded into the program and visualized. In an initial step (Fig. 4.36a), a number of tool-related characteristics need to be defined, including the tool symmetry (which determines whether the level-set function needs to be updated every step), and whether the shoulder is concave. The latter triggers the activation of the additional elements to fill the shoulder cavity associated with concave tools (see Sec. 4.3.5).

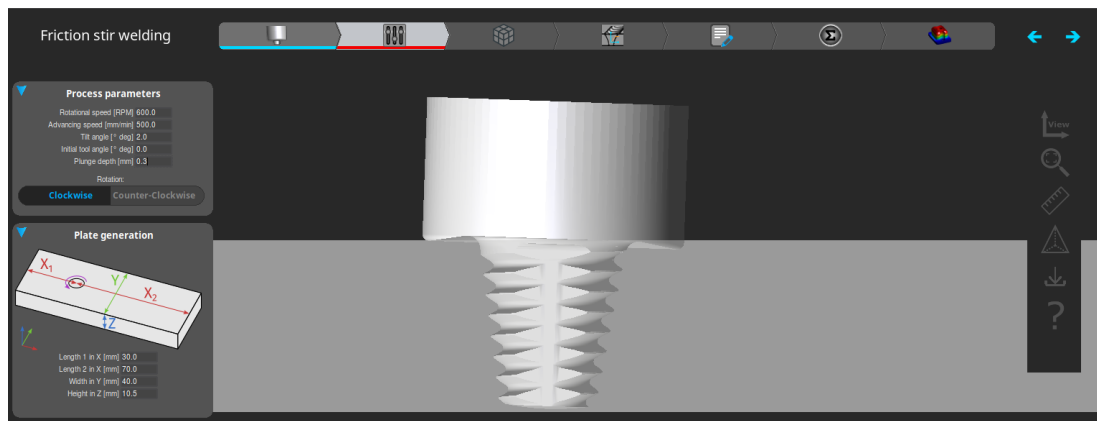
The computational domain is defined by its dimensions in x , y , and z , and by defining the tilt angle and plunging depth, the tool is positioned with respect to the computational domain (Fig. 4.36b).

Similarly, the remaining parameters such as friction and refinement parameters can be entered (see Fig. 4.36c), the temperature-dependent material properties can be modified, and the mesh is generated. Ultimately the GUI generates the input files which are required by the FEM kernel and launches the simulation – either locally or remotely on an HPC cluster.

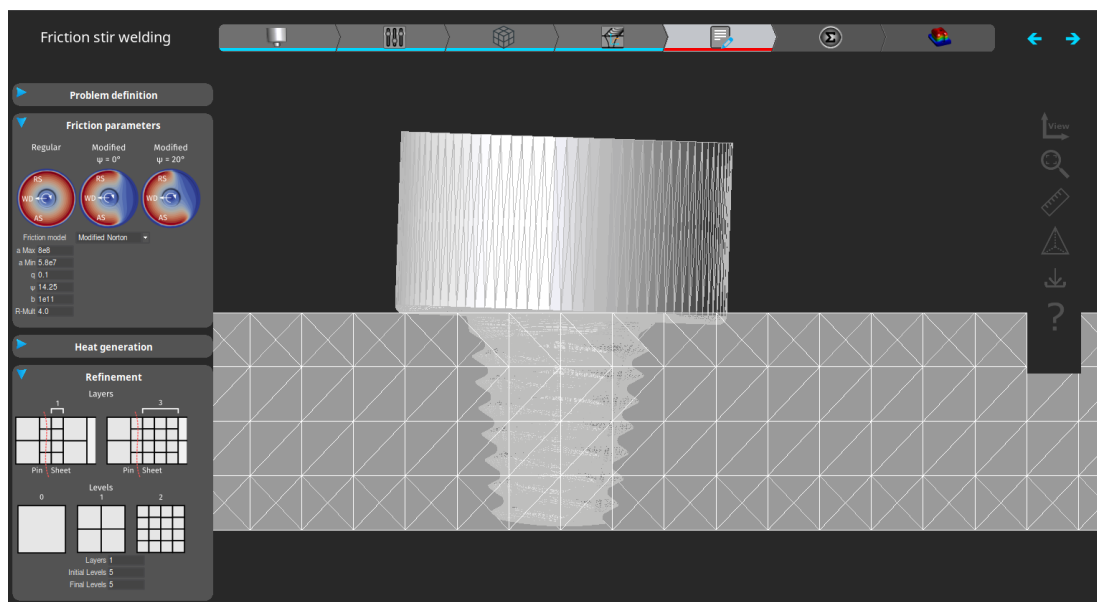
Table 4.1: List of most important FSW simulation parameters



(a) Tool characteristics



(b) Processing parameters and workpiece dimensions



(c) Friction coefficients and refinement strategy

Figure 4.36: Graphical User Interface

Postprocessing After completing the simulations, the relevant numerical results need to be gathered and visualized. A common technique to experimentally study heat generation and diffusion is to embed thermocouples in the workpiece and weld past them. The result is a graph showing the temperature evolution over time. An equivalent graph cannot be easily obtained from the simulation since it is performed in an Eulerian framework, where only the spacial temperature distribution is available. To obtain the temperature evolution of a material point, the temperature must be plotted along a streamline, as shown in Fig. 4.37.

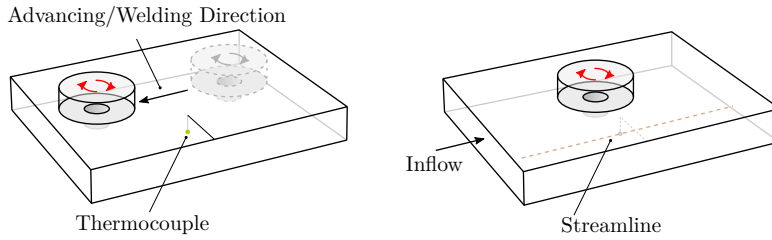


Figure 4.37: Temperature measurements in Lagrangian framework (material temperature, left) and Eulerian framework (spacial temperature, right)

Although the velocity field can be used to analyze the material flow, it is not a very intuitive representation of the material movement. This problem is shared by all fluids-based approaches that operate in an Eulerian framework (see Tab. 4.18). As a remedy, particle tracing can be utilized to better visualize the material flow.

A block of massless particles (tracers) is positioned inside the computational domain in front of the tool. Throughout the simulation, the tracer positions are advected based on the velocity field obtained by solving the Stokes flow problem. Dialami et al. [198] have shown that the time integration must be performed using a 4th-order Runge-Kutta method to achieve accurate results. Less accurate time integration methods, such as first order explicit time integration (explicit Euler), cannot preserve circular trajectories, leading to an un-physical radial spread.

A problem regarding particle tracing is the different timescales involved. Assuming a rotational velocity of 600 RPM, a full tool rotation takes 0.1 s. Assuming one rotation is discretized into 40 time steps, this results in a timestep size of 0.0025 s. Using the artificial speed-up parameter (see Eq. (4.3) in Sec. 4.3.2), a (periodic) steady-state solution can be reached in 2 to 3 rotations – or 80 to 120 time steps.

Much more time is needed for a particle to traverse the length of approximately three shoulder diameters (*i.e.*, from in front of the tool to the rear of the tool). Assuming an advancing speed of 0.5 m/min or 0.0083 m/s and an exemplary shoulder diameter of 15 mm, the time needed for the particle to pass the tool is 5.4 s or 2160 time steps.

The solution is clear for axisymmetric tools where a proper steady-state solution is reached. The particle tracing can be performed after finishing the thermo-mechanical simulations by simply using the final converged velocity field. For non-axisymmetric tools, a similar approach can be followed. After reaching the periodic steady state, the velocity fields of a full rotation are saved and periodically repeated to perform the particle tracing. Fig. 4.38 shows the results of a particle

[198]: Dialami et al. (2013), “Material flow visualization in Friction Stir Welding via particle tracing”

tracing simulation, where the temperature fields are mapped onto the the particles.

Implementation and High-Performance Computing (HPC)

All presented methods were implemented in FEMUSS (Finite Element Methods Using Subgrid Scales), an in-house Finite Element code developed at the International Center for Numerical Methods in Engineering (CIMNE). The code base is mostly written in Fortran 2003 and utilizes the full potential of an object-oriented programming language. FEMUSS is a multi-physics code capable of performing and coupling solid mechanics, compressible and incompressible fluid mechanics, convection-diffusion-reaction, as well as optics and acoustics simulations. As the name suggests, the software was developed around variational multi-scale stabilization. The software is fully parallelized using distributed memory parallelization [199] and supports efficient adaptive octree refinement and load-rebalancing [116].

Many simulations, including the mesh convergence studies with up to 50 million elements [195][Appx. C], were performed on LUMI, at the time of writing, Europe’s largest supercomputer.

[199]: MPI Forum (2015), *A Message-Passing Interface Standard Version 3.1*

[116]: Baiges et al. (2017), “Ref-efficientlib: An Efficient Load-Rebalanced Adaptive Mesh Refinement Algorithm for High-Performance Computational Physics Meshes”

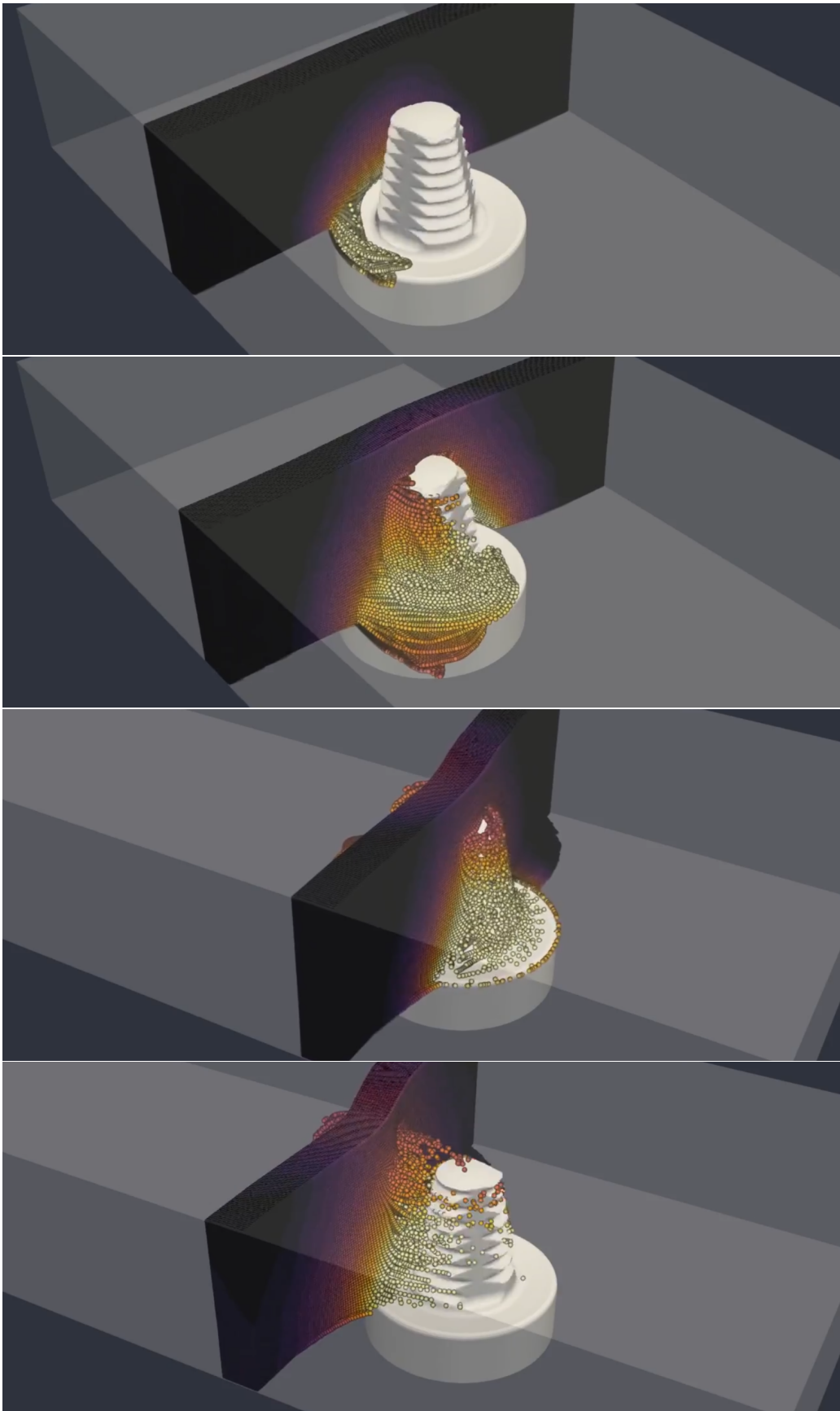


Figure 4.38: Particle Tracing

4.6 Concluding Remarks

The third part of this work addressed the application of methods to deal with isochoric material flow and non-body fitted methods to Friction Stir Welding – an emerging solid-state joining technique that is increasingly used for joining lightweight structural components for transportation applications (*i.e.*, automotive, rail, marine and aerospace sectors) and other load-bearing structures. The section addressed the fundamentals of FSW, the influence of tool shapes and process parameters, and the applicability to different materials. The advantages – most importantly being a solid-state joining process that operates at lower temperatures and avoids phase change – were outlined, and challenges were acknowledged. The most important one is its sensibility towards its processing parameters.

Naturally, numerical simulations are useful tools for engineers and researchers to study the effect of different parameters on the produced welds. Over the last two decades, increasingly complex approaches have been tried and tested to predict material flow, temperature evolution, and changes in microstructure, among other things. FSW simulations fall into two major categories: solids-based and fluids-based approaches. While the first is exclusively modeled in a Lagrangian framework, the second can be modeled in Eulerian or ALE frameworks. More recently, particle-based methods have been presented.

Fluids-based methods, which are particularly useful in studying the (periodic) steady-state solution of the actual welding phase (excluding plunging and extraction), only allow for unchanging meshes when assuming axisymmetric tools. Non-axisymmetric tools (which are the common case) are required for the ALE framework, and frequent remeshing and projections between meshes are mandatory. The Apropos Kinematic Framework was developed to address this issue but is limited to zero-degree tilt angles.

This work explores the possibilities that non-body-fitted methods offer. By describing the tool-workpiece interface using a discrete level-set function computed directly from a CAD file, arbitrary tool shapes and movements are possible, including the aforementioned tilt angles. Furthermore, preprocessing is reduced to a minimum since the work-intensive domain and mesh generation is greatly simplified. The use of unfitted meshes brings its own challenges, one being small cut elements and the other being the time dependency of the level-set function, which leads to the dynamic re-assignment of nodes and elements to the physical and non-physical subdomains. The first issue is addressed by using the Ghost Stabilization approach, which, due to its utilization of orthogonal projections, ties in well with the orthogonal subscales stabilization used for stabilizing the mixed velocity-pressure problem. The latter is addressed by computing a harmonic extension in the non-physical subdomain, ensuring optimal initial conditions for nodes that become part of the physical subdomain upon updating the level-set function due to the tool rotation.

Concave tools and the definition of surface normals presented additional challenges. To prevent unphysical contact between tools and workpieces, especially with concave shoulder features, an extra layer

of elements is generated above the physical domain. This layer fills the cavity formed by the concave shoulder. To achieve realistic surface normals independent of the level-set function's linear interpolation, the original normals are retrieved from the CAD file. The results align well with experimental data [195][Appx. C], and parameter studies demonstrate expected and justifiable responses in reaction forces, torque, and maximum temperature.

To bridge the gap between academic research and industrial application, significant effort was invested in creating pre- and post-processing tools. These tools streamline the generation of simulation input data and provide intuitive visualizations for the process analysis and the comparison with experimental data.

[195]: Venghaus et al. (2025), "Embedded technology for enhanced modeling of Friction Stir Welding processes"

5.1 Summary

Metal forming and manufacturing processes are a cornerstone of today's modern world. With the increasing complexity of manufacturing processes and the goal of using available resources as efficiently as possible, numerical simulations – particularly the Finite Element Method – have become essential tools for researchers and engineers. These tools can predict a wide range of physical quantities from displacements and stresses to temperatures and microstructures.

However, numerical analysis is a discipline in its own right, and the numerical modeling of mechanical processes is often not as straightforward as initially assumed. Numerical methods can become unstable, yielding unphysical or no results at all. This is especially true for metal forming applications involving (nearly) isochoric deformations, which occur during plastic deformations.

Also, while numerical tools are meant to save time compared to experiments, generating and handling geometries and meshes, defining simulation parameters, and applying boundary conditions can be time-consuming and require significant human interaction.

This work aimed to study existing methods to address these problems, evaluate their applicability to metal forming and manufacturing problems, and adapt and improve them where necessary.

Mixed finite element formulations have proven to be excellent tools for addressing problems in the isochoric domain (*i.e.*, mixed displacement-pressure formulations, u/p), as well as for ensuring local convergence of strains, even when using only linear elements (*i.e.*, displacement-strain formulation, u/ε). Additionally, a combination is possible (*i.e.*, displacement-pressure-deviatoric strain, $u/p/e$) that combines the benefits of both the u/p and u/ε formulations.

The downside of this *three-field* formulation is its large number of degrees of freedom per node. Part of this work was the development of the Adaptive Formulation Refinement to address this issue by selectively activating the enhanced three-field formulation based on physical criteria or error estimation. The conducted tests showed that, when using reasonable criteria, the results were on par with the three-field formulation. Further investigations revealed that special consideration must be given to the solver to fully capitalize on the gained efficiency. The method was then successfully applied to quasi-brittle structural failure analysis and combined with classical adaptive mesh (h -) refinement.

An important part of this work was enhancing the applicability of the developed methods to actual engineering problems, ensuring they did not remain an academic exercise. Friction Stir Welding (FSW) was

such a problem. Due to the enormous plastic and localized deformations, FSW is a prime candidate for using mixed formulations. Furthermore, the numerical modeling of FSW is characterized by complex geometries and moving boundary conditions.

This work used CutFEM to achieve three goals simultaneously. By allowing for non-body-fitted meshes, a purely Eulerian framework was used, eliminating any mesh movement and resulting re-meshing. By modeling the tool using a discrete level-set function computed from a regular CAD file, the pre-processing effort was minimized. And, ultimately, the tool could be of arbitrary shape and positioned at any arbitrary position and angle.

The obtained simulation results showed good agreement with experimental data, and the conducted parameter studies showcased the framework's ability to assist engineers in optimizing process parameters. To enhance the accessibility of the developed framework, a Graphical User Interface was developed to create simulation input files and to start and manage simulations. Furthermore, a particle tracing algorithm was implemented to provide an intuitive means of visualizing material flow.

5.2 Contribution

The main scientific contributions of the work are, for each of the parts of the thesis:

Part I *Element Technology*

- Implementation of the $u/p/e$ formulation and the respective stabilization algorithms in a fully parallelized, high-performance computing FE framework (FEMUSS).
- Conceptualization, development and implementation of the adaptive formulation refinement method.
- Adaptation of the OSS stabilization approach to deal with the $u/p-u/p/e$ interface in the context of Adaptive Formulation Refinement (AFR).
- Development of efficient, AFR specific, solver strategies to minimize simulation times.
- Adapting the AFR approach to be used in conjunction with adaptive mesh refinement.
- Application to fracture mechanics.

Results were published in the following scientific journals

- H. Venghaus, M. Chiumenti, J. Baiges, D. Juhre, and I. Castañar. “An accurate approach to simulate friction stir welding processes using adaptive formulation refinement”. In: *Finite Elements in Analysis and Design* 224 (2023) [Q1-ranked at time of publication] [Appx. A]
- G. Barbat, M. Cervera, H. Venghaus, and M. Chiumenti. “Objectivity in quasi-brittle structural failure via adaptive formulation

and mesh refinement”. In: *Theoretical and Applied Fracture Mechanics* 122 (2022)

[Q1-ranked at time of publication] [Appx. B]

and presented at the following scientific congresses

- H. Venghaus, M. Chiumenti, J. Baiges, and D. Juhre. “Adaptive Formulation Refinement Stress, strain and dissipation accurate FE technology for metal forming applications.” In: *XVI International Conference on Computational Plasticity. Fundamentals and Applications (COMPLAS)*. 2021
- H. Venghaus, G. Barbat, M. Chiumenti, J. Baiges, and D. Juhre. “Combined Refinement Strategies Using Mixed Finite Element Technology in Friction Stir Welding Analysis”. In: *9th GACM Colloquium on Computational Mechanics (GACM)*. 2022
- H. Venghaus, M. Chiumenti, J. Baiges, and D. Juhre. “Combined Refinement Strategies Using Mixed Finite Element Technology in Friction Stir Welding Analysis”. In: *Joint International Symposium on Friction Stir Welding and Processing (ISFSWP)*. 2022

Part II *Unfitted FEM / Embedded Domain Methods*

- Implementation of CAD-based discrete level-set functions to represent industrial geometries in an FEM framework. Access to CAD geometry is realized through the *OpenCascade* [74] library, requiring the implementation of a C++ FORTRAN interface. The implementation supports the following formats: STL, STEP, and IGES.
- Revision and partial re-implementation of boundary-conform integration algorithms for cut elements in FEMUSS
- Implementation of element decomposition strategy to apply boundary-conform integration to linear quadrilateral and hexahedral elements
- Implementation of regular and non-negative Moment Fitting algorithm
- Implementation of Penalty and Nitsche’s Method
- Implementation of Ghost Penalty stabilization for mechanical and thermal problems

[74]: (2024), *OpenCascade*

Since this section of the thesis largely deals with researching available methods, assessing their suitability, and implementing them, no publications were made as the work that was conducted would not have been considered a new contribution. The contributions are nevertheless listed here since they form the basis of the final chapter: the simulation of FSW using an embedded domain method.

Part III *Friction Stir Welding*

- Implementation of temperature-dependent linearized Norton-Hoff material law, as well as enhanced FSW friction laws, based off works by Dialami et al. [188]
- Development of a CAD-to-Simulation workflow for non-body fitted FSW simulations, including:

[188]: Dialami et al. (2017), “Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation”

- Conceptualization, development and implementation of automatic domain extension scheme to account for concave tools
- Implementation of harmonic extension for all unknowns
- Conceptualization, development and implementation of algorithm to recover original normals from CAD geometry using the previously established *OpenCascade* [74] interface
- Performance of parameter fitting and validation using experimental data
- Studying effects of different welding parameters such as advancing speed, rotation speed, plunging depth, tools scale and tilt angle on reaction forces, torque and temperature
- Enhancing accessibility and working towards facilitating technology transfer by developing a Graphical User Interface for easy generation of simulation input data
- Implementation of particle tracing algorithm based on works by Dialami et al. [198], with the addition of method to drastically reduce simulation time by “recycling” the velocity’s (periodic) steady-state solution

[74]: (2024), *OpenCascade*

[198]: Dialami et al. (2013), “Material flow visualization in Friction Stir Welding via particle tracing”

Results were published in the following scientific journals

- H. Venghaus, M. Chiumenti, J. Baiges, D. Juhre, and N. Dialami. “Embedded technology for enhanced modeling of Friction Stir Welding processes”. In: *Computer Methods in Applied Mechanics and Engineering* 435 (2025)
[Q1-ranked at time of publication] [Appx. C]

and presented at the following scientific congresses

- H. Venghaus, M. Chiumenti, J. Baiges, and D. Juhre. “A Comprehensive Framework for Accurate and Efficient Simulation of Friction Stir Welding Processes.” In: *XVII International Conference on Computational Plasticity. Fundamentals and Applications (COMPLAS)*. 2023

5.3 Future Work

This work lays a solid foundation for future research and opens up new possibilities to introduce advanced numerical methods to simulate metal forming and manufacturing processes.

Application to other industrial processes There are other metal forming and manufacturing processes which are also characterized by one or more of the following aspects: isochoric deformations, strain localizations, and complex and/or moving boundaries. All of them can benefit from adopting the presented numerical methods, namely mixed two- and three-field finite element formulations, stabilized with orthogonal sub-grid scales, while using CutFEM to allow for non-body fitted meshes and describing the boundary using a discrete level-set function computed from a CAD geometry. Possible applications include:

- Forging, where both die and workpiece could be described using a level-set function.

- Additive Manufacturing, especially for hybrid printing where the printed material is deposited onto existing components, *e.g.*, for printing additional features or to repair the component. Describing the existing components using a discrete level-set function computed from a CAD file can be very convenient.
- Extrusion, which exhibits temperatures and material flow at the die that are similar to FSW. The iterative processes of designing dies can be streamlined by using a discrete level-set function computed from a CAD file rather than creating body-fitted meshes each time the geometry is updated.

Further development of individual methods Different aspects of the developed and presented methods can be developed further.

- Study of stability of mixed three-field formulations in CutFEM and, if necessary, development of suitable Ghost stabilization terms.
- Application of Adaptive Formulation Refinement in the context of CutFEM.
- Development of fractional step schemes to enhance simulation times of mixed three-field formulations.

Further development of FSW framework Finally, the developed FSW framework can be adapted to enhance its accuracy further and to simulate more complicated scenarios.

- Application of Adaptive Formulation Refinement in the context of Embedded Friction Stir Welding. The enhanced strain accuracy of the three-field formulation will be beneficial to capture the strong gradients at the tool-workpiece interface.
- Introduction of further analysis tools to framework, such as multi-phase flow for dissimilar joining, metallurgical studies, or double-sided welding.

Bibliography

- [1] M. Cervera, M. Chiumenti, and C. Agelet de Saracibar. “Shear band localization via local J2 continuum damage mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* 193.9-11 (2004), pp. 849–880. DOI: [10.1016/j.cma.2003.11.009](https://doi.org/10.1016/j.cma.2003.11.009) (cit. on pp. 3, 20).
- [2] B Koohbor and S Serajzadeh. “Thermomechanical behaviours of strip and work-rolls in cold rolling process”. In: *The Journal of Strain Analysis for Engineering Design* 46.8 (2011), pp. 794–804. DOI: [10.1177/0309324711417693](https://doi.org/10.1177/0309324711417693) (cit. on p. 3).
- [3] European Commission. *Council backs Commission’s plans to invest €1 billion in world-class European supercomputers*. https://ec.europa.eu/commission/presscorner/detail/en/ip_18_5864 [Accessed: 2024-10-28]. 2018. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_18_5864 (cit. on p. 5).
- [4] Ronaldo I. Borja. *Plasticity*. Springer Berlin Heidelberg, 2013 (cit. on p. 10).
- [5] Thomas J.R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. New York: Dover Publications, 2000 (cit. on p. 11).
- [6] M. Cervera and M. Chiumenti. “Mesh objective tensile cracking via a local continuum damage model and a crack tracking technique”. In: *Computer Methods in Applied Mechanics and Engineering* 196.1-3 (2006), pp. 304–320. DOI: [10.1016/j.cma.2006.04.008](https://doi.org/10.1016/j.cma.2006.04.008) (cit. on p. 11).
- [7] Gabriel Barbat Vlad. “Mixed finite elements with independent strain interpolation for isotropic and orthotropic damage”. PhD thesis. Universitat Politècnica de Catalunya, 2021. DOI: [10.5821/dissertation-2117-346007](https://doi.org/10.5821/dissertation-2117-346007) (cit. on p. 11).
- [8] Savvas Saloustros, Miguel Cervera, and Luca Pelà. “Challenges, Tools and Applications of Tracking Algorithms in the Numerical Modelling of Cracks in Concrete and Masonry Structures”. In: *Archives of Computational Methods in Engineering* 26.4 (2018), pp. 961–1005. DOI: [10.1007/s11831-018-9274-3](https://doi.org/10.1007/s11831-018-9274-3) (cit. on p. 11).
- [9] O. C. Zienkiewicz, R. L. Taylor, and J. M. Too. “Reduced integration technique in general analysis of plates and shells”. In: *International Journal for Numerical Methods in Engineering* 3.2 (1971), pp. 275–290. DOI: [10.1002/nme.1620030211](https://doi.org/10.1002/nme.1620030211) (cit. on p. 12).
- [10] Isaac Fried. “Finite element analysis of incompressible material by residual energy balancing”. In: *International Journal of Solids and Structures* 10.9 (1974), pp. 993–1002. DOI: [10.1016/0020-7683\(74\)90007-9](https://doi.org/10.1016/0020-7683(74)90007-9) (cit. on p. 13).
- [11] D. J. Naylor. “Stresses in nearly incompressible materials by finite elements with application to the calculation of excess pore pressures”. In: *International Journal for Numerical Methods in Engineering* 8.3 (1974), pp. 443–460. DOI: [10.1002/nme.1620080302](https://doi.org/10.1002/nme.1620080302) (cit. on p. 13).
- [12] Thomas J. R. Hughes. “Generalization of selective integration procedures to anisotropic and nonlinear media”. In: *International Journal for Numerical Methods in Engineering* 15.9 (1980), pp. 1413–1418. DOI: [10.1002/nme.1620150914](https://doi.org/10.1002/nme.1620150914) (cit. on p. 14).
- [13] Thomas J. Hughes, Robert L. Taylor, and Jerome L. Sackman. *Finite Element Formulation and Solution of Contact-Impact Problems in Continuum Mechanics*. 1974 (cit. on p. 15).
- [14] J.C. Simo, R.L. Taylor, and K.S. Pister. “Variational and projection methods for the volume constraint in finite deformation elasto-plasticity”. In: *Computer Methods in Applied Mechanics and Engineering* 51.1-3 (1985), pp. 177–208. DOI: [10.1016/0045-7825\(85\)90033-7](https://doi.org/10.1016/0045-7825(85)90033-7) (cit. on p. 15).
- [15] Laurent Duchêne, A.B. Bettaieb, and Anne Habraken. “Assessment of the Enhanced Assumed Strain (EAS) and the Assumed Natural Strain (ANS) techniques in the mechanical behavior of the SSH3D Solid-Shell element”. In: *COMPLAS XI : proceedings of the XI International Conference on Computational Plasticity : fundamentals and applications*. <https://upcommons.upc.edu/handle/2117/184289>. 2011, pp. 1357–1368 (cit. on p. 15).

- [16] J. C. Simo and M. S. Rifai. “A class of mixed assumed strain methods and the method of incompatible modes”. In: *International Journal for Numerical Methods in Engineering* 29.8 (1990), pp. 1595–1638. DOI: [10.1002/nme.1620290802](https://doi.org/10.1002/nme.1620290802) (cit. on p. 15).
- [17] R.J. Alves de Sousa et al. “On the use of a reduced enhanced solid-shell (RESS) element for sheet forming simulations”. In: *International Journal of Plasticity* 23.3 (2007), pp. 490–515. DOI: [10.1016/j.ijplas.2006.06.004](https://doi.org/10.1016/j.ijplas.2006.06.004) (cit. on p. 16).
- [18] T. J. R. Hughes. “Equivalence of Finite Elements for Nearly Incompressible Elasticity”. In: *Journal of Applied Mechanics* 44.1 (1977), pp. 181–183. DOI: [10.1115/1.3423994](https://doi.org/10.1115/1.3423994) (cit. on p. 16).
- [19] M. Chiumenti et al. “A stabilized formulation for incompressible elasticity using linear displacement and pressure interpolations”. In: *Computer Methods in Applied Mechanics and Engineering* 191.46 (2002), pp. 5253–5264. DOI: [10.1016/s0045-7825\(02\)00443-7](https://doi.org/10.1016/s0045-7825(02)00443-7) (cit. on p. 20).
- [20] Nouredine Atalla, Raymond Panneton, and Patricia Debergue. “A mixed displacement-pressure formulation for poroelastic materials”. In: *The Journal of the Acoustical Society of America* 104.3 (1998), pp. 1444–1452. DOI: [10.1121/1.424355](https://doi.org/10.1121/1.424355) (cit. on p. 20).
- [21] M. Cervera et al. “Mixed linear/linear simplicial elements for incompressible elasticity and plasticity”. In: *Computer Methods in Applied Mechanics and Engineering* 192.49-50 (2003), pp. 5249–5263. DOI: [10.1016/j.cma.2003.07.007](https://doi.org/10.1016/j.cma.2003.07.007) (cit. on p. 20).
- [22] M. Cervera, M. Chiumenti, and C. Agelet de Saracibar. “Softening, localization and stabilization: capture of discontinuous solutions in J2 plasticity”. In: *International Journal for Numerical and Analytical Methods in Geomechanics* 28.5 (2004), pp. 373–393. DOI: [10.1002/nag.341](https://doi.org/10.1002/nag.341) (cit. on p. 20).
- [23] M. Cervera and M. Chiumenti. “Size effect and localization in J2 plasticity”. In: *International Journal of Solids and Structures* 46.17 (2009), pp. 3301–3312. DOI: [10.1016/j.ijsolstr.2009.04.025](https://doi.org/10.1016/j.ijsolstr.2009.04.025) (cit. on p. 20).
- [24] M. Cervera, M. Chiumenti, and R. Codina. “Mixed stabilized finite element methods in nonlinear solid mechanics. Part I: Formulation”. In: *Computer Methods in Applied Mechanics and Engineering* 199.37-40 (2010), pp. 2559–2570. DOI: [10.1016/j.cma.2010.04.006](https://doi.org/10.1016/j.cma.2010.04.006) (cit. on pp. 20, 22).
- [25] M. Cervera, M. Chiumenti, and R. Codina. “Mixed stabilized finite element methods in nonlinear solid mechanics. Part II: Strain localization”. In: *Computer Methods in Applied Mechanics and Engineering* 199.37-40 (2010), pp. 2571–2589. DOI: [10.1016/j.cma.2010.04.005](https://doi.org/10.1016/j.cma.2010.04.005) (cit. on p. 20).
- [26] Inocencio Castañar, Joan Baiges, and Ramon Codina. “A stabilized mixed finite element approximation for incompressible finite strain solid dynamics using a total Lagrangian formulation”. In: *Computer Methods in Applied Mechanics and Engineering* 368 (2020), p. 113164. DOI: [10.1016/j.cma.2020.113164](https://doi.org/10.1016/j.cma.2020.113164) (cit. on p. 20).
- [27] Inocencio Castañar et al. “Topological derivative-based topology optimization of incompressible structures using mixed formulations”. In: *Computer Methods in Applied Mechanics and Engineering* 390 (2022), p. 114438. DOI: [10.1016/j.cma.2021.114438](https://doi.org/10.1016/j.cma.2021.114438) (cit. on p. 20).
- [28] Inocencio Castañar, Joan Baiges, and Ramon Codina. “Topology optimization of incompressible structures subject to fluid–structure interaction”. In: *Structural and Multidisciplinary Optimization* 67.6 (2024). DOI: [10.1007/s00158-024-03770-6](https://doi.org/10.1007/s00158-024-03770-6) (cit. on p. 20).
- [29] Fucheng Tian et al. “Mixed displacement–pressure–phase field framework for finite strain fracture of nearly incompressible hyperelastic materials”. In: *Computer Methods in Applied Mechanics and Engineering* 394 (2022), p. 114933. DOI: [10.1016/j.cma.2022.114933](https://doi.org/10.1016/j.cma.2022.114933) (cit. on p. 20).
- [30] Chennakesava Kadapa and Mokarram Hossain. “A linearized consistent mixed displacement–pressure formulation for hyperelasticity”. In: *Mechanics of Advanced Materials and Structures* 29.2 (2020), pp. 267–284. DOI: [10.1080/15376494.2020.1762952](https://doi.org/10.1080/15376494.2020.1762952) (cit. on p. 20).
- [31] Ramon Codina, Inocencio Castañar, and Joan Baiges. “Finite element approximation of stabilized mixed models in finite strain hyperelasticity involving displacements and stresses and/or pressure—An overview of alternatives”. In: *International Journal for Numerical Methods in Engineering* 125.18 (2024). DOI: [10.1002/nme.7540](https://doi.org/10.1002/nme.7540) (cit. on pp. 20, 23).

- [32] M. Chiumenti et al. “Numerical modeling of friction stir welding processes”. In: *Computer Methods in Applied Mechanics and Engineering* 254 (2013), pp. 353–369. DOI: [10.1016/j.cma.2012.09.013](https://doi.org/10.1016/j.cma.2012.09.013) (cit. on pp. 20, 75).
- [33] C. Agelet de Saracibar et al. “Computational Modeling and Sub-Grid Scale Stabilization of Incompressibility and Convection in the Numerical Simulation of Friction Stir Welding Processes”. In: *Archives of Computational Methods in Engineering* 21.1 (2014), pp. 3–37. DOI: [10.1007/s11831-014-9094-z](https://doi.org/10.1007/s11831-014-9094-z) (cit. on pp. 20, 74).
- [34] Chennakesava Kadapa. “Mixed displacement–pressure formulations and suitable finite elements for multimaterial problems with compressible and incompressible models”. In: *Computer Methods in Applied Mechanics and Engineering* 432 (2024), p. 117354. DOI: [10.1016/j.cma.2024.117354](https://doi.org/10.1016/j.cma.2024.117354) (cit. on pp. 20, 21).
- [35] M. Chiumenti et al. “Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation”. In: *Finite Elements in Analysis and Design* 192 (2021), p. 103534. DOI: [10.1016/j.finel.2021.103534](https://doi.org/10.1016/j.finel.2021.103534) (cit. on pp. 20–22, 26).
- [36] G. Scovazzi, T. Song, and X. Zeng. “A velocity/stress mixed stabilized nodal finite element for elastodynamics: Analysis and computations with strongly and weakly enforced boundary conditions”. In: *Computer Methods in Applied Mechanics and Engineering* 325 (2017), pp. 532–576. DOI: [10.1016/j.cma.2017.07.018](https://doi.org/10.1016/j.cma.2017.07.018) (cit. on p. 22).
- [37] M. Chiumenti, M. Cervera, and R. Codina. “A mixed three-field FE formulation for stress accurate analysis including the incompressible limit”. In: *Computer Methods in Applied Mechanics and Engineering* 283 (2015), pp. 1095–1116. DOI: [10.1016/j.cma.2014.08.004](https://doi.org/10.1016/j.cma.2014.08.004) (cit. on p. 22).
- [38] Georgios Vlachakis et al. “Out-of-plane seismic response and failure mechanism of masonry structures using finite elements with enhanced strain accuracy”. In: *Engineering Failure Analysis* 97 (2019), pp. 534–555. DOI: [10.1016/j.engfailanal.2019.01.017](https://doi.org/10.1016/j.engfailanal.2019.01.017) (cit. on p. 23).
- [39] M. Cervera, G. B. Barbat, and M. Chiumenti. “Architecture of a multi-crack model with full closing, reopening and sliding capabilities”. In: *Computational Mechanics* 65.6 (2020), pp. 1593–1620. DOI: [10.1007/s00466-020-01836-y](https://doi.org/10.1007/s00466-020-01836-y) (cit. on p. 23).
- [40] G.B. Barbat et al. “Structural size effect: Experimental, theoretical and accurate computational assessment”. In: *Engineering Structures* 213 (2020), p. 110555. DOI: [10.1016/j.engstruct.2020.110555](https://doi.org/10.1016/j.engstruct.2020.110555) (cit. on p. 23).
- [41] Carlos A. Moreira et al. “Accurate thermal-induced structural failure analysis under incompressible conditions”. In: *Engineering Structures* 261 (2022), p. 114213. DOI: [10.1016/j.engstruct.2022.114213](https://doi.org/10.1016/j.engstruct.2022.114213) (cit. on p. 23).
- [42] Luan C. Trinh et al. “A strain-displacement mixed formulation based on the modified couple stress theory for the flexural behaviour of laminated beams”. In: *Composites Part B: Engineering* 185 (2020), p. 107740. DOI: [10.1016/j.compositesb.2019.107740](https://doi.org/10.1016/j.compositesb.2019.107740) (cit. on p. 23).
- [43] A. Aguirre, R. Codina, and J. Baiges. “Stress–displacement stabilized finite element analysis of thin structures using solid-shell elements, Part I: On the need of interpolating the stresses”. In: *Finite Elements in Analysis and Design* 236 (2024), p. 104168. DOI: [10.1016/j.finel.2024.104168](https://doi.org/10.1016/j.finel.2024.104168) (cit. on p. 23).
- [44] A. Aguirre et al. “Stress–displacement stabilized finite element analysis of thin structures using Solid-Shell elements, Part II: Finite strain hyperelasticity”. In: *Finite Elements in Analysis and Design* 236 (2024), p. 104179. DOI: [10.1016/j.finel.2024.104179](https://doi.org/10.1016/j.finel.2024.104179) (cit. on p. 23).
- [45] Laura Moreno et al. “Logarithmic conformation reformulation in viscoelastic flow problems approximated by a VMS-type stabilized finite element formulation”. In: *Computer Methods in Applied Mechanics and Engineering* 354 (2019), pp. 706–731. DOI: [10.1016/j.cma.2019.06.001](https://doi.org/10.1016/j.cma.2019.06.001) (cit. on p. 23).

- [46] Laura Moreno, Ramon Codina, and Joan Baiges. “Solution of transient viscoelastic flow problems approximated by a term-by-term VMS stabilized finite element formulation using time-dependent subgrid-scales”. In: *Computer Methods in Applied Mechanics and Engineering* 367 (2020), p. 113074. DOI: [10.1016/j.cma.2020.113074](https://doi.org/10.1016/j.cma.2020.113074) (cit. on p. 23).
- [47] Hiromichi Itou, Victor A. Kovtunenkov, and Evgeny M. Rudoy. “Three-field mixed formulation of elasticity model nonlinear in the mean normal stress for the problem of non-penetrating cracks in bodies”. In: *Applications in Engineering Science* 7 (2021), p. 100060. DOI: [10.1016/j.apples.2021.100060](https://doi.org/10.1016/j.apples.2021.100060) (cit. on p. 23).
- [48] Franco Brezzi and Michel Fortin, eds. *Mixed and Hybrid Finite Element Methods*. Springer New York, 1991 (cit. on p. 24).
- [49] C. Taylor and P. Hood. “A numerical solution of the Navier-Stokes equations using the finite element technique”. In: *Computers & Fluids* 1.1 (1973), pp. 73–100. DOI: [10.1016/0045-7930\(73\)90027-3](https://doi.org/10.1016/0045-7930(73)90027-3) (cit. on p. 24).
- [50] Jean Donea and Antonio Huerta. *Finite Element Methods for Flow Problems*. Wiley, 2003 (cit. on p. 24).
- [51] D. N. Arnold, F. Brezzi, and M. Fortin. “A stable finite element for the stokes equations”. In: *Calcolo* 21.4 (1984), pp. 337–344. DOI: [10.1007/bf02576171](https://doi.org/10.1007/bf02576171) (cit. on p. 24).
- [52] M. Crouzeix and P.-A. Raviart. “Conforming and nonconforming finite element methods for solving the stationary Stokes equations I”. In: *Revue française d’automatique informatique recherche opérationnelle. Mathématique* 7.R3 (1973), pp. 33–75. DOI: [10.1051/m2an/197307r300331](https://doi.org/10.1051/m2an/197307r300331) (cit. on p. 24).
- [53] Thomas J.R. Hughes. “Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods”. In: *Computer Methods in Applied Mechanics and Engineering* 127.1-4 (1995), pp. 387–401. DOI: [10.1016/0045-7825\(95\)00844-9](https://doi.org/10.1016/0045-7825(95)00844-9) (cit. on p. 25).
- [54] Thomas J.R. Hughes et al. “The variational multiscale method—a paradigm for computational mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* 166.1-2 (1998), pp. 3–24. DOI: [10.1016/s0045-7825\(98\)00079-6](https://doi.org/10.1016/s0045-7825(98)00079-6) (cit. on p. 25).
- [55] Ramon Codina. “On stabilized finite element methods for linear systems of convection–diffusion–reaction equations”. In: *Computer Methods in Applied Mechanics and Engineering* 188.1-3 (2000), pp. 61–82. DOI: [10.1016/s0045-7825\(00\)00177-8](https://doi.org/10.1016/s0045-7825(00)00177-8) (cit. on p. 25).
- [56] Ramon Codina. “Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods”. In: *Computer Methods in Applied Mechanics and Engineering* 190.13-14 (2000), pp. 1579–1599. DOI: [10.1016/s0045-7825\(00\)00254-1](https://doi.org/10.1016/s0045-7825(00)00254-1) (cit. on p. 25).
- [57] Santiago Badia and Ramon Codina. “Unified Stabilized Finite Element Formulations for the Stokes and the Darcy Problems”. In: *SIAM Journal on Numerical Analysis* 47.3 (2009), pp. 1971–2000. DOI: [10.1137/08072632x](https://doi.org/10.1137/08072632x) (cit. on pp. 25, 26).
- [58] Ramon Codina et al. “Variational Multiscale Methods in Computational Fluid Dynamics”. In: *Encyclopedia of Computational Mechanics Second Edition* (2017), pp. 1–28. DOI: [10.1002/9781119176817.ecm2117](https://doi.org/10.1002/9781119176817.ecm2117) (cit. on p. 27).
- [59] O. C. Zienkiewicz and J. Z. Zhu. “A simple error estimator and adaptive procedure for practical engineering analysis”. In: *International Journal for Numerical Methods in Engineering* 24.2 (1987), pp. 337–357. DOI: [10.1002/nme.1620240206](https://doi.org/10.1002/nme.1620240206) (cit. on p. 32).
- [60] Joan Baiges and Ramon Codina. “Variational Multiscale error estimators for solid mechanics adaptive simulations: An Orthogonal Subgrid Scale approach”. In: *Computer Methods in Applied Mechanics and Engineering* 325 (2017), pp. 37–55. DOI: [10.1016/j.cma.2017.07.008](https://doi.org/10.1016/j.cma.2017.07.008) (cit. on p. 33).
- [61] A. Bazile et al. “Variational Multiscale error estimator for anisotropic adaptive fluid mechanic simulations: Application to convection–diffusion problems”. In: *Computer Methods in Applied Mechanics and Engineering* 331 (2018), pp. 94–115. DOI: [10.1016/j.cma.2017.11.019](https://doi.org/10.1016/j.cma.2017.11.019) (cit. on p. 33).

- [62] Camilo Bayona-Roa, Ramon Codina, and Joan Baiges. “Variational multiscale error estimators for the adaptive mesh refinement of compressible flow simulations”. In: *Computer Methods in Applied Mechanics and Engineering* 337 (2018), pp. 501–526. DOI: [10.1016/j.cma.2018.03.010](https://doi.org/10.1016/j.cma.2018.03.010) (cit. on p. 33).
- [63] Satish Balay et al. *PETSc/TAO Users Manual*. Tech. rep. ANL-21/39 - Revision 3.16. Argonne National Laboratory, 2021 (cit. on p. 36).
- [64] H. Venghaus et al. “An accurate approach to simulate friction stir welding processes using adaptive formulation refinement”. In: *Finite Elements in Analysis and Design* 224 (2023), p. 103986. DOI: [10.1016/j.finel.2023.103986](https://doi.org/10.1016/j.finel.2023.103986) (cit. on pp. 36, 81, 99).
- [65] G. Barbat et al. “Objectivity in quasi-brittle structural failure via adaptive formulation and mesh refinement”. In: *Theoretical and Applied Fracture Mechanics* 122 (2022), p. 103646. DOI: [10.1016/j.tafmec.2022.103646](https://doi.org/10.1016/j.tafmec.2022.103646) (cit. on pp. 37, 99).
- [66] P. J. Green and R. Sibson. “Computing Dirichlet Tessellations in the Plane”. In: *The Computer Journal* 21.2 (1978), pp. 168–173. DOI: [10.1093/comjnl/21.2.168](https://doi.org/10.1093/comjnl/21.2.168) (cit. on p. 41).
- [67] Marc Vigo Anglada. “An improved incremental algorithm for constructing restricted Delaunay triangulations”. In: *Computers & Graphics* 21.2 (1997), pp. 215–223. DOI: [10.1016/s0097-8493\(96\)00085-4](https://doi.org/10.1016/s0097-8493(96)00085-4) (cit. on p. 41).
- [68] Hang Si. “TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator”. In: *ACM Transactions on Mathematical Software* 41.2 (2015), pp. 1–36. DOI: [10.1145/2629697](https://doi.org/10.1145/2629697) (cit. on p. 41).
- [69] Jason F. Shepherd and Chris R. Johnson. “Hexahedral mesh generation constraints”. In: *Engineering with Computers* 24.3 (2008), pp. 195–213. DOI: [10.1007/s00366-008-0091-4](https://doi.org/10.1007/s00366-008-0091-4) (cit. on p. 41).
- [70] James Gregson, Alla Sheffer, and Eugene Zhang. “All-Hex Mesh Generation via Volumetric Poly-Cube Deformation”. In: *Computer Graphics Forum* 30.5 (2011), pp. 1407–1416. DOI: [10.1111/j.1467-8659.2011.02015.x](https://doi.org/10.1111/j.1467-8659.2011.02015.x) (cit. on p. 41).
- [71] Manuel Meßmer et al. “Efficient CAD-integrated isogeometric analysis of trimmed solids”. In: *Computer Methods in Applied Mechanics and Engineering* 400 (2022), p. 115584. DOI: [10.1016/j.cma.2022.115584](https://doi.org/10.1016/j.cma.2022.115584) (cit. on p. 43).
- [72] N. Sukumar et al. “Modeling holes and inclusions by level sets in the extended finite-element method”. In: *Computer Methods in Applied Mechanics and Engineering* 190.46–47 (2001), pp. 6183–6200. DOI: [10.1016/s0045-7825\(01\)00215-8](https://doi.org/10.1016/s0045-7825(01)00215-8) (cit. on pp. 43, 44).
- [73] Frederic Gibou, Ronald Fedkiw, and Stanley Osher. “A review of level-set methods and some recent applications”. In: *Journal of Computational Physics* 353 (2018), pp. 82–109. DOI: [10.1016/j.jcp.2017.10.006](https://doi.org/10.1016/j.jcp.2017.10.006) (cit. on pp. 43, 44).
- [74] *OpenCascade*. 2024. URL: <https://dev.opencascade.org/> (visited on 06/25/2024) (cit. on pp. 45, 100, 101).
- [75] Kwok Wah Cheng and Thomas-Peter Fries. “Higher-order XFEM for curved strong and weak discontinuities”. In: *International Journal for Numerical Methods in Engineering* 82.5 (2009), pp. 564–590. DOI: [10.1002/nme.2768](https://doi.org/10.1002/nme.2768) (cit. on p. 46).
- [76] László Kudela et al. “Efficient and accurate numerical quadrature for immersed boundary methods”. In: *Advanced Modeling and Simulation in Engineering Sciences* 2.1 (2015). DOI: [10.1186/s40323-015-0031-y](https://doi.org/10.1186/s40323-015-0031-y) (cit. on p. 46).
- [77] Márton Pető et al. “Octree-based integration scheme with merged sub-cells for the finite cell method: Application to non-linear problems in 3D”. In: *Computer Methods in Applied Mechanics and Engineering* 401 (2022), p. 115565. DOI: [10.1016/j.cma.2022.115565](https://doi.org/10.1016/j.cma.2022.115565) (cit. on p. 47).
- [78] Márton Pető et al. “Boolean finite cell method for multi-material problems including local enrichment of the Ansatz space”. In: *Computational Mechanics* 72.4 (2023), pp. 743–764. DOI: [10.1007/s00466-023-02305-y](https://doi.org/10.1007/s00466-023-02305-y) (cit. on pp. 47, 54).

- [79] A. Düster et al. “The finite cell method for three-dimensional problems of solid mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* 197.45–48 (2008), pp. 3768–3782. DOI: [10.1016/j.cma.2008.02.036](https://doi.org/10.1016/j.cma.2008.02.036) (cit. on pp. 47, 50, 54).
- [80] László Kudela et al. “Smart octrees: Accurately integrating discontinuous functions in 3D”. In: *Computer Methods in Applied Mechanics and Engineering* 306 (2016), pp. 406–426. DOI: [10.1016/j.cma.2016.04.006](https://doi.org/10.1016/j.cma.2016.04.006) (cit. on p. 47).
- [81] Simeon Hubrich et al. “Numerical integration of discontinuous functions: moment fitting and smart octree”. In: *Computational Mechanics* 60.5 (2017), pp. 863–881. DOI: [10.1007/s00466-017-1441-0](https://doi.org/10.1007/s00466-017-1441-0) (cit. on p. 47).
- [82] B. Müller, F. Kummer, and M. Oberlack. “Highly accurate surface and volume integration on implicit domains by means of moment-fitting”. In: *International Journal for Numerical Methods in Engineering* 96.8 (2013), pp. 512–528. DOI: [10.1002/nme.4569](https://doi.org/10.1002/nme.4569) (cit. on p. 47).
- [83] Wadhah Garhuom and Alexander Düster. “Non-negative moment fitting quadrature for cut finite elements and cells undergoing large deformations”. In: *Computational Mechanics* 70.5 (2022), pp. 1059–1081. DOI: [10.1007/s00466-022-02203-9](https://doi.org/10.1007/s00466-022-02203-9) (cit. on p. 48).
- [84] Ivo Babuska. “The Finite Element Method with Penalty”. In: *Mathematics of Computation* 27.122 (1973), p. 221. DOI: [10.2307/2005611](https://doi.org/10.2307/2005611) (cit. on p. 49).
- [85] Ivo Babuška and Miloš Zlámal. “Nonconforming Elements in the Finite Element Method with Penalty”. In: *SIAM Journal on Numerical Analysis* 10.5 (1973), pp. 863–875. DOI: [10.1137/0710071](https://doi.org/10.1137/0710071) (cit. on p. 49).
- [86] J. Nitsche. “Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind”. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 36.1 (1971), pp. 9–15. DOI: [10.1007/bf02995904](https://doi.org/10.1007/bf02995904) (cit. on p. 49).
- [87] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*. Springer New York, 2004 (cit. on pp. 49, 50).
- [88] R. Glowinski et al. “A distributed Lagrange multiplier/fictitious domain method for particulate flows”. In: *International Journal of Multiphase Flow* 25.5 (1999), pp. 755–794. DOI: [10.1016/s0301-9322\(98\)00048-2](https://doi.org/10.1016/s0301-9322(98)00048-2) (cit. on p. 50).
- [89] Erik Burman. “Ghost penalty”. In: *Comptes Rendus. Mathématique* 348.21–22 (2010), pp. 1217–1220. DOI: [10.1016/j.crma.2010.10.006](https://doi.org/10.1016/j.crma.2010.10.006) (cit. on pp. 50, 51).
- [90] Erik Burman et al. “CutFEM: Discretizing geometry and partial differential equations”. In: *International Journal for Numerical Methods in Engineering* 104.7 (2014), pp. 472–501. DOI: [10.1002/nme.4823](https://doi.org/10.1002/nme.4823) (cit. on p. 50).
- [91] Monique Dauge, Alexander Düster, and Ernst Rank. “Theoretical and Numerical Investigation of the Finite Cell Method”. In: *Journal of Scientific Computing* 65.3 (2015), pp. 1039–1064. DOI: [10.1007/s10915-015-9997-3](https://doi.org/10.1007/s10915-015-9997-3) (cit. on p. 50).
- [92] Wadhah Garhuom, Khuldoon Usman, and Alexander Düster. “An eigenvalue stabilization technique to increase the robustness of the finite cell method for finite strain problems”. In: *Computational Mechanics* 69.5 (2022), pp. 1225–1240. DOI: [10.1007/s00466-022-02140-7](https://doi.org/10.1007/s00466-022-02140-7) (cit. on p. 51).
- [93] Stefan Loehnert. “A stabilization technique for the regularization of nearly singular extended finite elements”. In: *Computational Mechanics* 54.2 (2014), pp. 523–533. DOI: [10.1007/s00466-014-1003-7](https://doi.org/10.1007/s00466-014-1003-7) (cit. on p. 51).
- [94] Stefan Loehnert. “Stabilizing the XFEM for static and dynamic crack simulations”. In: *PAMM* 15.1 (2015), pp. 137–138. DOI: [10.1002/pamm.201510059](https://doi.org/10.1002/pamm.201510059) (cit. on p. 51).
- [95] Nicolas Moës, John Dolbow, and Ted Belytschko. “A finite element method for crack growth without remeshing”. In: *International Journal for Numerical Methods in Engineering* 46.1 (1999), pp. 131–150. DOI: [10.1002/\(sici\)1097-0207\(19990910\)46:1<131::aid-nme726>3.0.co;2-j](https://doi.org/10.1002/(sici)1097-0207(19990910)46:1<131::aid-nme726>3.0.co;2-j) (cit. on p. 54).

- [96] N. Sukumar et al. “Extended finite element method for three-dimensional crack modelling”. In: *International Journal for Numerical Methods in Engineering* 48.11 (2000), pp. 1549–1570. DOI: [10.1002/1097-0207\(20000820\)48:11<1549::aid-nme955>3.0.co;2-a](https://doi.org/10.1002/1097-0207(20000820)48:11<1549::aid-nme955>3.0.co;2-a) (cit. on p. 54).
- [97] T. Strouboulis, I. Babuška, and K. Copps. “The design and analysis of the Generalized Finite Element Method”. In: *Computer Methods in Applied Mechanics and Engineering* 181.1–3 (2000), pp. 43–69. DOI: [10.1016/S0045-7825\(99\)00072-9](https://doi.org/10.1016/S0045-7825(99)00072-9) (cit. on p. 54).
- [98] T. Strouboulis, K. Copps, and I. Babuška. “The generalized finite element method: an example of its implementation and illustration of its performance”. In: *International Journal for Numerical Methods in Engineering* 47.8 (2000), pp. 1401–1417. DOI: [10.1002/\(sici\)1097-0207\(20000320\)47:8<1401::aid-nme835>3.0.co;2-8](https://doi.org/10.1002/(sici)1097-0207(20000320)47:8<1401::aid-nme835>3.0.co;2-8) (cit. on p. 54).
- [99] J.M. Melenk and I. Babuška. “The partition of unity finite element method: Basic theory and applications”. In: *Computer Methods in Applied Mechanics and Engineering* 139.1–4 (1996), pp. 289–314. DOI: [10.1016/S0045-7825\(96\)01087-0](https://doi.org/10.1016/S0045-7825(96)01087-0) (cit. on p. 54).
- [100] I. Babuška and J. M. Melenk. “The Partition of Unity Method”. In: *International Journal for Numerical Methods in Engineering* 40.4 (1997), pp. 727–758. DOI: [10.1002/\(sici\)1097-0207\(19970228\)40:4<727::aid-nme86>3.0.co;2-n](https://doi.org/10.1002/(sici)1097-0207(19970228)40:4<727::aid-nme86>3.0.co;2-n) (cit. on p. 54).
- [101] Michael Griebel and Marc Alexander Schweitzer. “A Particle-Partition of Unity Method for the Solution of Elliptic, Parabolic, and Hyperbolic PDEs”. In: *SIAM Journal on Scientific Computing* 22.3 (2000), pp. 853–890. DOI: [10.1137/S1064827599355840](https://doi.org/10.1137/S1064827599355840) (cit. on p. 54).
- [102] Thomas-Peter Fries and Ted Belytschko. “The extended/generalized finite element method: An overview of the method and its applications”. In: *International Journal for Numerical Methods in Engineering* 84.3 (2010), pp. 253–304. DOI: [10.1002/nme.2914](https://doi.org/10.1002/nme.2914) (cit. on p. 54).
- [103] A M Roma, C S Peskin, and M J Berger. “An adaptive version of the Immersed Boundary Method”. In: *Journal of Computational Physics* 153.2 (1999). DOI: [10.1006/jcph.1999.6293](https://doi.org/10.1006/jcph.1999.6293) (cit. on p. 54).
- [104] Charles S. Peskin. “The immersed boundary method”. In: *Acta Numerica* 11 (2002), pp. 479–517. DOI: [10.1017/S0962492902000077](https://doi.org/10.1017/S0962492902000077) (cit. on p. 54).
- [105] Anita Hansbo and Peter Hansbo. “An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems”. In: *Computer Methods in Applied Mechanics and Engineering* 191.47–48 (2002), pp. 5537–5552. DOI: [10.1016/S0045-7825\(02\)00524-8](https://doi.org/10.1016/S0045-7825(02)00524-8) (cit. on p. 54).
- [106] Isabelle Ramière, Philippe Angot, and Michel Belliard. “A fictitious domain approach with spread interface for elliptic problems with general boundary conditions”. In: *Computer Methods in Applied Mechanics and Engineering* 196.4–6 (2007), pp. 766–781. DOI: [10.1016/j.cma.2006.05.012](https://doi.org/10.1016/j.cma.2006.05.012) (cit. on p. 54).
- [107] Jamshid Parvizian, Alexander Düster, and Ernst Rank. “Finite cell method”. In: *Computational Mechanics* 41.1 (2007), pp. 121–133. DOI: [10.1007/s00466-007-0173-y](https://doi.org/10.1007/s00466-007-0173-y) (cit. on p. 54).
- [108] Erik Burman and Peter Hansbo. “Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method”. In: *Computer Methods in Applied Mechanics and Engineering* 199.41–44 (2010), pp. 2680–2686. DOI: [10.1016/j.cma.2010.05.011](https://doi.org/10.1016/j.cma.2010.05.011) (cit. on p. 54).
- [109] Erik Burman and Peter Hansbo. “Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method”. In: *Applied Numerical Mathematics* 62.4 (2012), pp. 328–341. DOI: [10.1016/j.apnum.2011.01.008](https://doi.org/10.1016/j.apnum.2011.01.008) (cit. on p. 54).
- [110] Erik Burman and Peter Hansbo. “Fictitious domain methods using cut elements: III. A stabilized Nitsche method for Stokes’ problem”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 48.3 (2014), pp. 859–874. DOI: [10.1051/m2an/2013123](https://doi.org/10.1051/m2an/2013123) (cit. on p. 54).
- [111] S. Duzcek et al. “Numerical analysis of Lamb waves using the finite and spectral cell methods: LAMB WAVE ANALYSIS USING THE FCM AND THE SCM”. In: *International Journal for Numerical Methods in Engineering* 99.1 (2014), pp. 26–53. DOI: [10.1002/nme.4663](https://doi.org/10.1002/nme.4663) (cit. on p. 54).
- [112] Anthony T Patera. “A spectral element method for fluid dynamics: Laminar flow in a channel expansion”. In: *Journal of Computational Physics* 54.3 (1984), pp. 468–488. DOI: [10.1016/0021-9991\(84\)90128-1](https://doi.org/10.1016/0021-9991(84)90128-1) (cit. on pp. 54, 55).

- [113] A. Main and G. Scovazzi. “The shifted boundary method for embedded domain computations. Part I: Poisson and Stokes problems”. In: *Journal of Computational Physics* 372 (2018), pp. 972–995. DOI: [10.1016/j.jcp.2017.10.026](https://doi.org/10.1016/j.jcp.2017.10.026) (cit. on pp. 54, 55).
- [114] A. Main and G. Scovazzi. “The shifted boundary method for embedded domain computations. Part II: Linear advection–diffusion and incompressible Navier–Stokes equations”. In: *Journal of Computational Physics* 372 (2018), pp. 996–1026. DOI: [10.1016/j.jcp.2018.01.023](https://doi.org/10.1016/j.jcp.2018.01.023) (cit. on pp. 54, 55).
- [115] Joan Baiges et al. “An adaptive Fixed-Mesh ALE method for free surface flows”. In: *Computer Methods in Applied Mechanics and Engineering* 313 (2017), pp. 159–188. DOI: [10.1016/j.cma.2016.09.041](https://doi.org/10.1016/j.cma.2016.09.041) (cit. on pp. 55, 56).
- [116] Joan Baiges and Camilo Bayona. “Refficientlib: An Efficient Load-Rebalanced Adaptive Mesh Refinement Algorithm for High-Performance Computational Physics Meshes”. In: *SIAM Journal on Scientific Computing* 39.2 (2017), pp. C65–C95. DOI: [10.1137/15m105330x](https://doi.org/10.1137/15m105330x) (cit. on pp. 56, 94).
- [117] Wayne M. Thomas et al. *FRICTION WELDING*. PCT/GB92/02203. 1995 (cit. on p. 60).
- [118] Narges Dialami. “Thermo-mechanical analysis of welding processes”. PhD thesis. Universitat Politècnica de Catalunya. Departament de Resistència de Materials i Estructures a l’Enginyeria, 2014 (cit. on pp. 60, 62, 63, 65, 66).
- [119] Suleyman Kilic, Fahrettin Ozturk, and Mehmet Fatih Demirdogen. “A comprehensive literature review on friction stir welding: Process parameters, joint integrity, and mechanical properties”. In: *Journal of Engineering Research* (2023). DOI: [10.1016/j.jer.2023.09.005](https://doi.org/10.1016/j.jer.2023.09.005) (cit. on pp. 60–63, 65, 67, 68, 70).
- [120] Pratik H Shah and Vishvesh J Badheka. “Friction stir welding of aluminium alloys: An overview of experimental findings – Process, variables, development and applications”. In: *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications* 233.6 (2017), pp. 1191–1226. DOI: [10.1177/1464420716689588](https://doi.org/10.1177/1464420716689588) (cit. on p. 61).
- [121] A. Bastier et al. “Modelling of the residual state of friction stir welded plates”. In: *Journal of Materials Processing Technology* 200.1–3 (2008), pp. 25–37. DOI: [10.1016/j.jmatprotec.2007.10.083](https://doi.org/10.1016/j.jmatprotec.2007.10.083) (cit. on p. 61).
- [122] Debasish Mishra et al. “A review on sensor based monitoring and control of friction stir welding process and a roadmap to Industry 4.0”. In: *Journal of Manufacturing Processes* 36 (2018), pp. 373–397. DOI: [10.1016/j.jmapro.2018.10.016](https://doi.org/10.1016/j.jmapro.2018.10.016) (cit. on p. 61).
- [123] Devang Sejani, Wenya Li, and Vivek Patel. “Stationary shoulder friction stir welding – low heat input joining technique: a review in comparison with conventional FSW and bobbin tool FSW”. In: *Critical Reviews in Solid State and Materials Sciences* 47.6 (2021), pp. 865–914. DOI: [10.1080/10408436.2021.1935724](https://doi.org/10.1080/10408436.2021.1935724) (cit. on p. 61).
- [124] Pabitra Maji et al. “An overview on friction stir welding/processing tools”. In: *Materials Today: Proceedings* 58 (2022), pp. 57–64. DOI: [10.1016/j.matpr.2022.01.009](https://doi.org/10.1016/j.matpr.2022.01.009) (cit. on pp. 63, 64).
- [125] R.J. Ding. “Retractable Pin-Tool Technology for Friction Stir Welding”. English. In: Cited by: 3; Conference name: Trends in Welding Research: Proceedings of the 5th International Conference; Conference date: 1 June 1998 through 5 June 1998; Conference code: 62380. 1998, pp. 585–589 (cit. on p. 63).
- [126] F. Marie, D. Alléhaux, and B. Esmiller. “Development of the Bobbin Tool Technique on Various Al-Alloys”. In: *5th Symposium on FSW*. Metz, France, 2005 (cit. on p. 63).
- [127] Robert Kosturek et al. “The influence of tool rotation speed on the low-cycle fatigue behavior of AA2519-T62 friction stir welded butt joints”. In: *Engineering Failure Analysis* 142 (2022), p. 106756. DOI: [10.1016/j.engfailanal.2022.106756](https://doi.org/10.1016/j.engfailanal.2022.106756) (cit. on p. 65).
- [128] AluStir. *FSW: Tool Tilt - How far Should the Tool be Tilted during Friction Stir Welding?* <https://www.alustir.com/english/fsw-tool-tilt/> [Accessed: 2024-10-15]. — (cit. on p. 65).

- [129] ESAB. *Friction Stir Welding - the ESAB Way*. <http://eng.esab.co.kr/Web-App/Upload/2012/04/05/FrictionStirWelding.pdf>. 2004. URL: <http://eng.esab.co.kr/Web-App/Upload/2012/04/05/Friction%20Stir%20Welding.pdf> (cit. on p. 65).
- [130] S. Verma, Meenu, and J.P. Misra. “Study on temperature distribution during Friction Stir Welding of 6082 aluminum alloy”. In: *Materials Today: Proceedings* 4.2 (2017), pp. 1350–1356. DOI: [10.1016/j.matpr.2017.01.156](https://doi.org/10.1016/j.matpr.2017.01.156) (cit. on p. 66).
- [131] Ashish M. Desai et al. “Friction stir welding of AZ31 magnesium alloy: A review”. In: *Materials Today: Proceedings* 47 (2021), pp. 6576–6584. DOI: [10.1016/j.matpr.2021.03.082](https://doi.org/10.1016/j.matpr.2021.03.082) (cit. on p. 66).
- [132] Kulwant Singh, Gurbhinder Singh, and Harmeet Singh. “Review on friction stir welding of magnesium alloys”. In: *Journal of Magnesium and Alloys* 6.4 (2018), pp. 399–416. DOI: [10.1016/j.jma.2018.06.001](https://doi.org/10.1016/j.jma.2018.06.001) (cit. on p. 66).
- [133] Răducu Nicolae Bulacu et al. “A Review on Friction Stir Welding of Copper: Tool Geometry, Process Parameters, and Joint Properties”. In: *Materials* 17.21 (2024), p. 5374. DOI: [10.3390/ma17215374](https://doi.org/10.3390/ma17215374) (cit. on p. 66).
- [134] R.S. Mishra and Z.Y. Ma. “Friction stir welding and processing”. In: *Materials Science and Engineering: R: Reports* 50.1–2 (2005), pp. 1–78. DOI: [10.1016/j.mser.2005.07.001](https://doi.org/10.1016/j.mser.2005.07.001) (cit. on pp. 66, 67).
- [135] R Nandan, T Debroy, and H Bhadeshia. “Recent advances in friction-stir welding – Process, weldment structure and properties”. In: *Progress in Materials Science* 53.6 (2008), pp. 980–1023. DOI: [10.1016/j.pmatsci.2008.05.001](https://doi.org/10.1016/j.pmatsci.2008.05.001) (cit. on p. 66).
- [136] G. Çam. “Friction stir welded structural materials: beyond Al-alloys”. In: *International Materials Reviews* 56.1 (2011), pp. 1–48. DOI: [10.1179/095066010x12777205875750](https://doi.org/10.1179/095066010x12777205875750) (cit. on p. 66).
- [137] Isak Nielsen. “Modeling and Control of Friction Stir Welding in 5 cm thick Copper Canisters”. MA thesis. Automatic Control, 2012, p. 103 (cit. on pp. 66, 70).
- [138] Akbar Heidarzadeh et al. “Development of a fuzzy logic based model to elucidate the effect of FSW parameters on the ultimate tensile strength and elongation of pure copper joints”. In: *Journal of Manufacturing Processes* 53 (2020), pp. 250–259. DOI: [10.1016/j.jmapro.2020.02.020](https://doi.org/10.1016/j.jmapro.2020.02.020) (cit. on p. 66).
- [139] Dhanesh G Mohan and ChuanSong Wu. “A Review on Friction Stir Welding of Steels”. In: *Chinese Journal of Mechanical Engineering* 34.1 (2021). DOI: [10.1186/s10033-021-00655-3](https://doi.org/10.1186/s10033-021-00655-3) (cit. on p. 66).
- [140] Ana I. Álvarez et al. “Assisted Friction Stir Welding of Carbon Steel: Use of Induction and Laser as Preheating Techniques”. In: *Friction Stir Welding and Processing VII*. Springer International Publishing, 2013, pp. 117–126. DOI: [10.1007/978-3-319-48108-1_13](https://doi.org/10.1007/978-3-319-48108-1_13) (cit. on p. 66).
- [141] Massab Junaid et al. “Comparison of microstructure, mechanical properties, and residual stresses in tungsten inert gas, laser, and electron beam welding of Ti-5Al-2.5Sn titanium alloy”. In: *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications* 233.7 (2017), pp. 1336–1351. DOI: [10.1177/1464420717748345](https://doi.org/10.1177/1464420717748345) (cit. on p. 67).
- [142] Sergei Tarasov et al. “Friction Stir Welding of Ti-6Al-4V Using a Liquid-Cooled Nickel Superalloy Tool”. In: *Technologies* 10.6 (2022), p. 118. DOI: [10.3390/technologies10060118](https://doi.org/10.3390/technologies10060118) (cit. on p. 67).
- [143] Meghnath Sen, Sachindra Shankar, and Somnath Chattopadhyaya. “Investigations into FSW joints of dissimilar aluminum alloys”. In: *Materials Today: Proceedings* 27 (2020), pp. 2455–2462. DOI: [10.1016/j.matpr.2019.09.218](https://doi.org/10.1016/j.matpr.2019.09.218) (cit. on p. 67).
- [144] Sachindra Shankar et al. “Dissimilar friction stir welding of Al to non-Al metallic materials: An overview”. In: *Materials Chemistry and Physics* 288 (2022), p. 126371. DOI: [10.1016/j.matchemphys.2022.126371](https://doi.org/10.1016/j.matchemphys.2022.126371) (cit. on p. 67).
- [145] P. G. Kossakowski, W. Weiślik, and M. Bakalarz. “Effect of Selected Friction Stir Welding Parameters on Mechanical Properties of Joints”. In: *Archives of Civil Engineering* 65.4 (2019), pp. 51–62. DOI: [10.2478/ace-2019-0046](https://doi.org/10.2478/ace-2019-0046) (cit. on p. 67).

- [146] M. W. Mahoney et al. “Properties of friction-stir-welded 7075 T651 aluminum”. In: *Metallurgical and Materials Transactions A* 29.7 (1998), pp. 1955–1964. DOI: [10.1007/s11661-998-0021-5](https://doi.org/10.1007/s11661-998-0021-5) (cit. on p. 67).
- [147] Yutaka S. Sato and Hiroyuki Kokawa. “Distribution of tensile property and microstructure in friction stir weld of 6063 aluminum”. In: *Metallurgical and Materials Transactions A* 32.12 (2001), pp. 3023–3031. DOI: [10.1007/s11661-001-0177-8](https://doi.org/10.1007/s11661-001-0177-8) (cit. on p. 67).
- [148] Hongjun Li, Jian Gao, and Qinchuan Li. “Fatigue of Friction Stir Welded Aluminum Alloy Joints: A Review”. In: *Applied Sciences* 8.12 (2018), p. 2626. DOI: [10.3390/app8122626](https://doi.org/10.3390/app8122626) (cit. on p. 67).
- [149] S.W. Williams and A. Steuwer. “Residual stresses in friction stir welding”. In: *Friction Stir Welding*. Elsevier, 2010, pp. 215–244. DOI: [10.1533/9781845697716.2.215](https://doi.org/10.1533/9781845697716.2.215) (cit. on p. 68).
- [150] J.O. Almen and P.H. Black. *Residual Stresses and Fatigue in Metals*. McGraw-Hill, 1963 (cit. on p. 68).
- [151] Tanveer Majeed et al. “Friction stir welding: A sustainable manufacturing process”. In: *Materials Today: Proceedings* 46 (2021), pp. 6558–6563. DOI: [10.1016/j.matpr.2021.04.025](https://doi.org/10.1016/j.matpr.2021.04.025) (cit. on p. 68).
- [152] Amber Shrivastava, Manuela Krones, and Frank E. Pfefferkorn. “Comparison of energy consumption and environmental impact of friction stir welding and gas metal arc welding for aluminum”. In: *CIRP Journal of Manufacturing Science and Technology* 9 (2015), pp. 159–168. DOI: [10.1016/j.cirpj.2014.10.001](https://doi.org/10.1016/j.cirpj.2014.10.001) (cit. on p. 68).
- [153] Hasan I. Dawood, Kahtan S. Mohammed, and Mumtaz Y. Rajab. “Advantages of the Green Solid State FSW over the Conventional GMAW Process”. In: *Advances in Materials Science and Engineering* 2014 (2014), pp. 1–10. DOI: [10.1155/2014/105713](https://doi.org/10.1155/2014/105713) (cit. on p. 68).
- [154] Guoqing Wang, Yanhua Zhao, and Yunfei Hao. “Friction stir welding of high-strength aerospace aluminum alloy and application in rocket tank manufacturing”. In: *Journal of Materials Science & Technology* 34.1 (2018), pp. 73–91. DOI: [10.1016/j.jmst.2017.11.041](https://doi.org/10.1016/j.jmst.2017.11.041) (cit. on p. 69).
- [155] Koki Nimura et al. “Technologies and Prospects of the H-IIB Launch Vehicle”. In: *Mitsubishi Heavy Industries Technical Review* 50.1 (2013). <https://www.mhi.co.jp/technology/review/pdf/e501/e501063.pdf>, pp. 63–70 (cit. on p. 69).
- [156] AluStir. *FSW: Rolling Stock*. <https://www.alustir.com/english/fsw-rolling-stock/> [Accessed: 2025-03-11]. — (cit. on pp. 69, 70).
- [157] TWI. *Railway Rolling Stock Manufacturers Implement Friction Stir Welding*. <https://www.twi-global.com/technical-knowledge/published-papers/railway-rolling-stock-manufacturers-implement-friction-stir-welding-october-2002> [Accessed: 2025-03-11]. 2002 (cit. on pp. 69, 70).
- [158] Ralph Gay, Mark Robinson, and Donato Zangani. “Crashworthiness of Joints in Aluminium Rail Vehicles”. In: (2014). https://www.researchgate.net/publication/265021938_Crashworthiness_of_Joints_in_Aluminium_Rail_Vehicles [Accessed: 2025-03-11] (cit. on pp. 69, 70).
- [159] Elizabeth Hoyos et al. “Case Study: Implementation of FSW in the Colombian Rail Transport Sector”. In: *Friction Stir Welding and Processing XI*. Springer International Publishing, 2021, pp. 3–12. DOI: [10.1007/978-3-030-65265-4_1](https://doi.org/10.1007/978-3-030-65265-4_1) (cit. on pp. 69, 70).
- [160] MachineDesign. *Car bodies first to be friction stir welded*. <https://www.machinedesign.com/archive/article/21816493/car-bodies-first-to-be-friction-stir-welded> [Accessed: 2025-03-11]. 2003 (cit. on p. 70).
- [161] TWI. *Friction Stir Welding in Series Production*. <https://www.twi-global.com/technical-knowledge/published-papers/friction-stir-welding-in-series-production-october-2004> [Accessed: 2025-03-11]. 2004 (cit. on p. 70).
- [162] TWI. *Friction Welded Wheels provide improved performance and reduced running costs*. <https://web.archive.org/web/20100815024249/http://www.twi.co.uk/content/autonews.html> [Accessed: 2025-03-11]. — (cit. on p. 70).

- [163] AppleInsider. *Apple slims down iMac 40% with 'friction-stir welding' & ditching the disc drive*. <https://appleinsider.com/articles/12/10/24/apple-slims-down-imac-40-with-friction-stir-welding-ditching-the-disc-drive> [Accessed: 2025-03-11]. 2012 (cit. on p. 70).
- [164] Wei Tang et al. "Friction Stir Welding – an Advanced Approach to Repair Nuclear Power Plant Components". In: *Proceedings of the 20th International Conference on Environmental Degradation of Materials in Nuclear Power Systems - Water Reactors (EnvDeg 2022)*. <https://www.ornl.gov/publication/friction-stir-welding-advanced-approach-repair-nuclear-power-plant-components>. Snowmass Village, Colorado, United States of America: AMPP, 2022, p. 17314 (cit. on p. 70).
- [165] Yuh J. Chao, X. Qi, and W. Tang. "Heat Transfer in Friction Stir Welding—Experimental and Numerical Studies". In: *Journal of Manufacturing Science and Engineering* 125.1 (2003), pp. 138–145. DOI: [10.1115/1.1537741](https://doi.org/10.1115/1.1537741) (cit. on p. 71).
- [166] C Chen and R Kovacevic. "Thermomechanical modelling and force analysis of friction stir welding by the finite element method". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 218.5 (2004), pp. 509–519. DOI: [10.1243/095440604323052292](https://doi.org/10.1243/095440604323052292) (cit. on p. 71).
- [167] J. J. Monaghan. "Smoothed Particle Hydrodynamics". In: *Annual Review of Astronomy and Astrophysics* 30.1 (1992), pp. 543–574. DOI: [10.1146/annurev.aa.30.090192.002551](https://doi.org/10.1146/annurev.aa.30.090192.002551) (cit. on p. 71).
- [168] M. B. Liu and G. R. Liu. "Smoothed Particle Hydrodynamics (SPH): an Overview and Recent Developments". In: *Archives of Computational Methods in Engineering* 17.1 (2010), pp. 25–76. DOI: [10.1007/s11831-010-9040-7](https://doi.org/10.1007/s11831-010-9040-7) (cit. on p. 71).
- [169] S. G. Psakhie et al. "Method of movable cellular automata as a tool for simulation within the framework of mesomechanics". In: *Russian Physics Journal* 38.11 (1995), pp. 1157–1168. DOI: [10.1007/bf00559396](https://doi.org/10.1007/bf00559396) (cit. on p. 71).
- [170] Wenxiao Pan et al. "A new smoothed particle hydrodynamics non-Newtonian model for friction stir welding: Process modeling and simulation of microstructure evolution in a magnesium alloy". In: *International Journal of Plasticity* 48 (2013), pp. 189–204. DOI: [10.1016/j.ijplas.2013.02.013](https://doi.org/10.1016/j.ijplas.2013.02.013) (cit. on p. 71).
- [171] Hamed Jamshidi Aval. "Smoothed-Particle Hydrodynamics (SPH) Simulation of AA6061-AA5086 Dissimilar Friction Stir Welding". In: *Metals* 13.5 (2023), p. 906. DOI: [10.3390/met13050906](https://doi.org/10.3390/met13050906) (cit. on p. 71).
- [172] Elizaveta Shishova et al. "Reversible inter-particle bonding in SPH for improved simulation of friction stir welding". In: *Computational Particle Mechanics* 10.3 (2022), pp. 555–564. DOI: [10.1007/s40571-022-00510-9](https://doi.org/10.1007/s40571-022-00510-9) (cit. on p. 71).
- [173] A. Yu. Smolin et al. "Understanding the mechanisms of friction stir welding based on computer simulation using particles". In: *Defence Technology* 14.6 (2018), pp. 643–656. DOI: [10.1016/j.dt.2018.09.003](https://doi.org/10.1016/j.dt.2018.09.003) (cit. on p. 71).
- [174] Simon Guerdoux et al. "Numerical simulation of the friction stir welding process using both Lagrangian and Arbitrary Lagrangian Eulerian formulations". In: *Materials Processing and Design: Modeling, Simulation and Applications - NUMIFORM 2004 - Proceedings of the 8th International Conference on Numerical Methods in Industrial Forming Processes*. Ed. by S Ghosh, Castro JM, and Lee JK. Vol. 712. <http://link.aip.org/link/?APCPCS/712/1259/1>. Columbus, United States: American Institute of Physics, 2004, ISBN : 0-7354-0188-8–pages 1259–1264. DOI: [10.1063/1.1766702](https://doi.org/10.1063/1.1766702) (cit. on p. 73).
- [175] Philippe Bussetta et al. "Comparison of a Fluid and a Solid Approach for the Numerical Simulation of Friction Stir Welding with a Non-Cylindrical Pin". In: *steel research international* 85.6 (2013), pp. 968–979. DOI: [10.1002/srin.201300182](https://doi.org/10.1002/srin.201300182) (cit. on p. 73).
- [176] T. Sheppard and D. S. Wright. "Determination of flow stress: Part 1 constitutive equation for aluminium alloys at elevated temperatures". In: *Metals Technology* 6.1 (1979), pp. 215–223. DOI: [10.1179/030716979803276264](https://doi.org/10.1179/030716979803276264) (cit. on p. 73).

- [177] Daniel Dall'Onder dos Santos et al. "Numerical approximations for flow of viscoplastic fluids in a lid-driven cavity". In: *Journal of Non-Newtonian Fluid Mechanics* 166.12–13 (2011), pp. 667–679. DOI: [10.1016/j.jnnfm.2011.03.004](https://doi.org/10.1016/j.jnnfm.2011.03.004) (cit. on p. 73).
- [178] Pierre J. Carreau. "Rheological Equations from Molecular Network Theories". In: *Transactions of the Society of Rheology* 16.1 (1972), pp. 99–127. DOI: [10.1122/1.549276](https://doi.org/10.1122/1.549276) (cit. on p. 73).
- [179] F.H. Norton. *The Creep of Steel at High Temperatures*. Bulletin. McGraw-Hill book Company, Incorporated, 1929 (cit. on pp. 73, 80).
- [180] Nicholas John Hoff. "Approximate analysis of structures in the presence of moderately large creep deformations". In: *Quarterly of Applied Mathematics* 12 (1954), pp. 49–55 (cit. on pp. 73, 80).
- [181] T. De Vuyst and L. D'Alvise. "Material Flow Around a Friction Stir Welding Tool: Development of a Thermo-Fluid Code". In: *Welding in the World* 51.3–4 (2007), pp. 37–43. DOI: [10.1007/bf03266558](https://doi.org/10.1007/bf03266558) (cit. on p. 74).
- [182] Mohamed Assidi and Lionel Fourment. "Accurate 3D friction stir welding simulation tool based on friction model calibration". In: *12th ESAFORM Conference on Material Forming*. Ed. by A.H. edited by van den Boogaard and University Twente of R. Akkermann. Vol. 2. Supplement 1. Enschede, Netherlands: Springer Verlag, 2009, Pages 327–330. DOI: [10.1007/s12289-009-0541-6](https://doi.org/10.1007/s12289-009-0541-6) (cit. on p. 74).
- [183] C. Agelet de Saracibar et al. "A Computational Model for the Numerical Simulation of FSW Processes". In: *AIP Conference Proceedings*. AIP, 2010, pp. 81–88. DOI: [10.1063/1.3457640](https://doi.org/10.1063/1.3457640) (cit. on p. 74).
- [184] Simon Guerdoux and Lionel Fourment. "A 3D numerical simulation of different phases of friction stir welding". In: *Modelling and Simulation in Materials Science and Engineering* 17.7 (2009), Article Number 075001 (32 p.) DOI: [10.1088/0965-0393/17/7/075001](https://doi.org/10.1088/0965-0393/17/7/075001) (cit. on p. 75).
- [185] Dongjoon Myung et al. "Probing the Mechanism of Friction Stir Welding with ALE Based Finite Element Simulations and Its Application to Strength Prediction of Welded Aluminum". In: *Metals and Materials International* 27.4 (2020), pp. 650–666. DOI: [10.1007/s12540-020-00901-8](https://doi.org/10.1007/s12540-020-00901-8) (cit. on p. 75).
- [186] Narges Dialami et al. "An apropos kinematic framework for the numerical modeling of friction stir welding". In: *Computers & Structures* 117 (2013), pp. 48–57. DOI: [10.1016/j.compstruc.2012.12.006](https://doi.org/10.1016/j.compstruc.2012.12.006) (cit. on p. 75).
- [187] Narges Dialami, Miguel Cervera, and Michele Chiumenti. "Effect of the Tool Tilt Angle on the Heat Generation and the Material Flow in Friction Stir Welding". In: *Metals* 9.1 (2018), p. 28. DOI: [10.3390/met9010028](https://doi.org/10.3390/met9010028) (cit. on pp. 76, 78).
- [188] N. Dialami et al. "Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation". In: *International Journal of Mechanical Sciences* 133 (2017), pp. 555–567. DOI: [10.1016/j.ijmecsci.2017.09.022](https://doi.org/10.1016/j.ijmecsci.2017.09.022) (cit. on pp. 76, 78, 83, 90, 100).
- [189] N. Dialami et al. "A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding". In: *International Journal of Mechanical Sciences* 122 (2017), pp. 215–227. DOI: [10.1016/j.ijmecsci.2016.12.016](https://doi.org/10.1016/j.ijmecsci.2016.12.016) (cit. on pp. 76, 81, 84).
- [190] Narges Dialami et al. "Experimental Validation of an FSW Model with an Enhanced Friction Law: Application to a Threaded Cylindrical Pin Tool". In: *Metals* 7.11 (2017), p. 491. DOI: [10.3390/met7110491](https://doi.org/10.3390/met7110491) (cit. on pp. 76, 78).
- [191] E. Feulvarch, J.-C. Roux, and J.-M. Bergheau. "A simple and robust moving mesh technique for the finite element simulation of Friction Stir Welding". In: *Journal of Computational and Applied Mathematics* 246 (2013), pp. 269–277. DOI: [10.1016/j.cam.2012.07.013](https://doi.org/10.1016/j.cam.2012.07.013) (cit. on p. 76).
- [192] Hugo Robe et al. "R-ALE simulation of heat transfer during friction stir welding of an AA2xxx/AA7xxx joint on a large process window". In: *International Journal of Mechanical Sciences* 155 (2019), pp. 31–40. DOI: [10.1016/j.ijmecsci.2019.02.029](https://doi.org/10.1016/j.ijmecsci.2019.02.029) (cit. on p. 76).

- [193] Philippe Bussetta et al. “Two 3D thermomechanical numerical models of friction stir welding processes with a trigonal pin”. In: *Numerical Heat Transfer, Part A: Applications* 70.9 (2016), pp. 995–1008. DOI: [10.1080/10407782.2016.1214498](https://doi.org/10.1080/10407782.2016.1214498) (cit. on p. 76).
- [194] Uwe Janoske, Markus Bürger, and Tobias Geiger. “Modelling of friction stir welding processes with a coupled immersed boundary-/volume of fluid-approach”. In: *Progress in Computational Fluid Dynamics, An International Journal* 19.1 (2019), p. 55. DOI: [10.1504/pcfd.2019.097602](https://doi.org/10.1504/pcfd.2019.097602) (cit. on pp. 77, 78).
- [195] H. Venghaus et al. “Embedded technology for enhanced modeling of Friction Stir Welding processes”. In: *Computer Methods in Applied Mechanics and Engineering* 435 (2025), p. 117539. DOI: [10.1016/j.cma.2024.117539](https://doi.org/10.1016/j.cma.2024.117539) (cit. on pp. 79, 83, 89, 90, 94, 97, 101).
- [196] N. Dialami et al. “Challenges in Thermo-mechanical Analysis of Friction Stir Welding Processes”. In: *Archives of Computational Methods in Engineering* 24.1 (2016), pp. 189–225. DOI: [10.1007/s11831-015-9163-y](https://doi.org/10.1007/s11831-015-9163-y) (cit. on p. 83).
- [197] Ramon Codina et al. “The fixed-mesh ALE approach for the numerical approximation of flows in moving domains”. In: *Journal of Computational Physics* 228.5 (2009), pp. 1591–1611. DOI: [10.1016/j.jcp.2008.11.004](https://doi.org/10.1016/j.jcp.2008.11.004) (cit. on p. 85).
- [198] N. Dialami et al. “Material flow visualization in Friction Stir Welding via particle tracing”. In: *International Journal of Material Forming* 8.2 (2013), pp. 167–181. DOI: [10.1007/s12289-013-1157-4](https://doi.org/10.1007/s12289-013-1157-4) (cit. on pp. 93, 101).
- [199] MPI Forum. *A Message-Passing Interface Standard Version 3.1*. Tech. rep. MPI Forum, 2015 (cit. on p. 94).
- [200] H. Venghaus et al. “Adaptive Formulation Refinement Stress, strain and dissipation accurate FE technology for metal forming applications.” In: *XVI International Conference on Computational Plasticity. Fundamentals and Applications (COMPLAS)*. 2021 (cit. on p. 100).
- [201] H. Venghaus et al. “Combined Refinement Strategies Using Mixed Finite Element Technology in Friction Stir Welding Analysis”. In: *9th GACM Colloquium on Computational Mechanics (GACM)*. 2022 (cit. on p. 100).
- [202] H. Venghaus et al. “Combined Refinement Strategies Using Mixed Finite Element Technology in Friction Stir Welding Analysis”. In: *Joint International Symposium on Friction Stir Welding and Processing (ISFSWP)*. 2022 (cit. on p. 100).
- [203] H. Venghaus et al. “A Comprehensive Framework for Accurate and Efficient Simulation of Friction Stir Welding Processes.” In: *XVII International Conference on Computational Plasticity. Fundamentals and Applications (COMPLAS)*. 2023 (cit. on p. 101).

Image Sources / Links

- Fig. 1.1a** https://www.rail-wheels.com/photo/rail-wheels/editor/20220111134652_56372.jpg
- Fig. 1.1b** <https://sourceint.com/wp-content/uploads/Sl-Aluminum-Extrusion-2-1-1.jpg>
- Fig. 1.1c** https://metallonline.ch/fileadmin/_processed_/7/3/csm_m_07_24_gruene_schienen_04_720e35bce2.jpeg
- Fig. 1.2a** <https://twi-global.com/image-library/hero/dual-rotation-friction-stir-welding.jpg>
- Fig. 1.2b** <https://us.arrk.com/wp-content/uploads/2023/06/3d-printing-how-does-direct-metal-laser-sintering-work-768x384.webp>
- Fig. 1.2c** <https://pick3dprinter.com/wp-content/uploads/2021/05/Directed-Energy-Deposition-3D-Printing.jpg>
- Fig. 2.1a** <https://www.radheshyaminvestmentcasting.com/wp-content/uploads/2018/06/abt-img.jpg>
- Fig. 2.1b** https://upload.wikimedia.org/wikipedia/commons/thumb/b/b2/Landslide_in_Cusco%2C_Peru_-_2018.jpg/800px-Landslide_in_Cusco%2C_Peru_-_2018.jpg
- Fig. 2.1c** <https://www.bridgebearing.org/img/lead-rubber-bearing-bridges.jpg>
- Fig. 2.1d** https://image.freepik.com/free-photo/process-bending-sheet-metal-hydraulic-bending-machine-metalworking-plant_347372-228.jpg
- Fig. 4.4a** https://bil-ibs.be/sites/default/files/lasprocessen/fsw/fsw_machine.jpg
- Fig. 4.4b** https://www.t-welding.com/images_2021/Friction%20Stir%20Welding%20Handbook.pdf
- Fig. 4.4c** https://www.kuka.com/-/media/kuka-corporate/images/industries/case-studies/fsw-portugal/fsw_portugal_welding_friction-stir-welding.jpg?rev=-1&w=767&hash=A32D0B027A02B6D330FD944B8075DD08
- Fig. 4.9a** https://en.m.wikipedia.org/wiki/File:SpaceX_factory_Falcon_9_booster_tank.jpg
- Fig. 4.9b** <https://www.nasa.gov/image-article/welding-underway-orion-first-artemis-mission-landing-astronauts-moon/>
- Fig. 4.9c** <https://image.jimcdn.com/app/cms/image/transf/dimension=624x10000:format=png/path/seae8da23f0f34385/image/id717e6ed3cf3c46f/version/1614758436/seitenwand-einer-bei-bombardier-aus-fsw-paneelen-hergestellten-zugs-der-londoner-u-bahn.png>
- Fig. 4.9d** <https://image.jimcdn.com/app/cms/image/transf/dimension=301x10000:format=jpg/path/seae8da23f0f34385/image/i7a8c1fa1727d24c7/version/1614758453/am-twi-ausgestelltes-teil-einer-bei-bombardier-aus-fsw-paneelen-hergestellten-londoner-u-bahn.jpg>

Publication 1

A



Contents lists available at ScienceDirect

Finite Elements in Analysis & Design

journal homepage: www.elsevier.com/locate/finel



An accurate approach to simulate friction stir welding processes using adaptive formulation refinement

H. Venghaus^{a,*}, M. Chiumenti^a, J. Baiges^{b,a}, D. Juhre^c, I. Castañar^b

^a International Center for Numerical Methods in Engineering (CIMNE), Polytechnic University of Catalonia (UPC), Campus Norte UPC, 08034 Barcelona, Spain

^b Polytechnic University of Catalonia (UPC), Barcelona Tech, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain

^c Institute of Mechanics (IFME), Otto von Guericke University Magdeburg (OVGU), Universitätsplatz 2, 39106 Magdeburg, Germany

ARTICLE INFO

Keywords:

Friction Stir Welding
Finite element method
Mixed formulations
Adaptive refinement

ABSTRACT

A novel *Adaptive Formulation Refinement* (AFR) strategy for Friction Stir Welding (FSW) problems is presented. In FSW, the accurate computation of strains is crucial to correctly predict the highly non-linear material behavior in the stir zone. Based on a posteriori error estimation, AFR switches between two mixed formulations depending on the required accuracy in the different regions of the domain. The higher accuracy formulation is used in the thermo-mechanically affected zone (TMAZ), while a computationally cheaper formulation is used elsewhere. AFR adds to the well-known h - (mesh size), p - (polynomial degree) and r -refinement (spatial distribution) approaches. The considered mixed formulations are the velocity/pressure ($\mathbf{u}/p$) and the velocity/pressure/deviatoric strain rate ($\mathbf{u}/p/e$) formulations—both suitable for isochoric material flow. By applying the AFR strategy, the use of linear elements is preserved, the incompressible flow of the material is captured correctly and any remeshing is avoided. Furthermore, the treatment of the interface between refined and unrefined subdomains is straightforward due to the compatibility of variable fields and lack of hanging nodes. The accuracy of the results obtained from the AFR method compares favorably with reference results of the non-adaptive $\mathbf{u}/p/e$ formulation. At the same time, faster build and solve times are achieved.

1. Introduction

Friction Stir Welding (FSW) is a solid-state joining process, where two facing work pieces are welded together using a non-consumable pin-tool which rotates and moves along their interface. In FSW, the joining process is achieved because of heat generated by friction between the pin-tool shoulder and the work pieces as well as by the plastic deformation of the stirred material. Typically, the material is not melted. This is a great advantage over conventional welding techniques, since both solidification and liquation cracking are avoided. However, the FSW process is very sensitive to the process parameters such as the rotation speed, the advancing speed, the pin shape and tilting, among others. The numerical simulation of FSW is a valuable tool to understand the influence of these parameters to optimize the process [1].

The main goals of the simulation of FSW are: the prediction of the temperature distribution [2,3], the sensitivity to the process parameters [4], the analysis of the strain rate field [5] and the material flow [6,7] as well as the contact and friction [8,9]. The choice of the most suitable computational frameworks (i.e., Eulerian, Lagrangian or ALE) has also been analyzed [10]. Furthermore,

* Corresponding author.

E-mail address: hvenghaus@cimne.upc.edu (H. Venghaus).

<https://doi.org/10.1016/j.finel.2023.103986>

Received 21 November 2022; Received in revised form 28 March 2023; Accepted 24 May 2023

Available online 7 July 2023

0168-874X/© 2023 Elsevier B.V. All rights reserved.

several works are devoted to the evolution of the metallurgy in the Heat Affected Zone (HAZ) [11] and the formation of different defects induced by the welding process [6].

The material characterization in FSW simulations is also of great scientific interest because of the strong thermo-mechanical coupling. On one hand, the thermal analysis is driven by the heat generated by the plastic dissipation (material stirring and friction between tool and work piece). On the other hand, the thermal softening, induced by the temperature field, must be accounted for, in order to characterize both material stiffness and strength accurately. Moreover, the isochoric nature of the plastic deformations (i.e., J_2 -plasticity or visco-plastic flow models) must be considered [12]. As a consequence, the use of standard (displacement-based) formulations is typically precluded and more sophisticated Finite Element technologies are mandatory in order to deal with the inherent material incompressibility.

Established methods to deal with isochoric behavior are the *Reduced and Selective Integration* method [13–15], *B-bar* and *F-bar* approaches [16–19], the *Enhanced Assumed Strain* methods [20–22] or the *Q1P0 Formulation* [23]. Unfortunately these methods are often only applicable to nearly incompressible cases and/or are not compatible with arbitrary types of elements and choices of shape functions. Another approach to deal with incompressible behavior is the use of Discontinuous Galerkin methods [24–26].

The mixed velocity/pressure ($\mathbf{u}/p$) formulation is an alternative approach to deal with a fully isochoric behavior, regardless of the element type (e.g., hexahedral, tetrahedral or prism elements are allowed) and thus facilitating the mesh generation when complex industrial tools are used. Stabilization techniques and, particularly, the Variational Multi-Scale (VMS) method [27,28] have been introduced to overcome the limitation on the compatibility among the interpolation degrees of the nodal variables. This approach, originally developed for incompressible computational fluid dynamics (CFD) analysis, has been extended to solid mechanics [29] to deal with elasto-plasticity, damage and strain localization problems [30–35], hyper-elasticity [36,37] and elasto-dynamics, including finite strain solid dynamics [38–41]. Throughout the present paper the $\mathbf{u}/p$ formulation will also be referred to as the *two-field* formulation.

To enhance the accuracy of the solution of a Finite Element simulation, the most frequently applied techniques are: the use of finer meshes (*h*-refinement), higher order shape functions (*p*-refinement) or a combination of both (*hp*-refinement [42,43]). Unfortunately, *p*-refinement is not straightforward for the $\mathbf{u}/p$ formulation. However, mixed formulations, considering either the stress or strain as an additional primary unknown, have proven to be a good alternative to the use of higher order shape functions.

The mixed (linear/linear) displacement/stress ($\mathbf{u}/\sigma$) formulation [44,45] was developed to enhance the solution accuracy by considering the stresses, additionally to the displacements, as primary unknowns. By splitting the stress tensor into its volumetric (i.e., the pressure p) and its deviatoric part ($\mathbf{s}$), it is possible to define a three-field $\mathbf{u}/p/\mathbf{s}$ formulation [46] able to combine the material incompressibility requirement with stress accuracy.

However, several material laws are strain driven. Therefore, the mixed displacement/strain ($\mathbf{u}/\epsilon$) formulation [44,47,48] was proposed. Finally, the novel displacement/pressure/deviatoric strain ($\mathbf{u}/p/\epsilon$) formulation [12] combines both the $\mathbf{u}/p$ and $\mathbf{u}/\epsilon$ formulations. The $\mathbf{u}/p/\epsilon$ formulation yields a faster convergence rate upon mesh refinement when compared to the classical $\mathbf{u}/p$ formulation and shows an enhanced accuracy for the same number of degrees of freedom. It has also proven to be a valuable tool in the analysis of non-linear, inelastic behavior [49] and topology optimization [50]. Throughout the present paper the $\mathbf{u}/p/\epsilon$ formulation will also be referred to as the *three-field* formulation.

When simulating FSW processes however, the solution is smooth in most regions of the computational domain. Therefore the use of the very sophisticated and CPU-expensive $\mathbf{u}/p/\epsilon$ formulation is unnecessary in most parts of the domain. Only close to the pin, where high temperature gradients and strain rates are observed, the $\mathbf{u}/p/\epsilon$ formulation is required. In order to define the most suitable and computationally efficient framework for the FSW simulation, an adaptive use of both two- and three-field formulations is presented in this work. This Adaptive Formulation Refinement (AFR) approach is driven by an a posteriori error estimator [51–54]. Thus, it is possible to switch between the $\mathbf{u}/p$ and $\mathbf{u}/p/\epsilon$ formulation according to the required solution accuracy, while preserving the use of linear elements and avoiding any remeshing. Furthermore, it can be shown that the treatment of the interface between unrefined ($\mathbf{u}/p$) and refined ($\mathbf{u}/p/\epsilon$) subdomains is straightforward, due to the compatibility of the variable fields and the lack of hanging nodes.

The main novelty of the present paper is the element-wise use of differently accurate (and computationally demanding) mixed and stabilized FE formulations as a means of adaptive refinement strategy. The presented paper shows that the superior strain rate accuracy of the three-field formulation can be applied in crucial areas of the computational domain while the overall computational cost is reduced by using the two-field formulation elsewhere. A reasonable subdivision of the computational domain into $\mathbf{u}/p$ and $\mathbf{u}/p/\epsilon$ elements is key to minimizing the global error with respect to the computational cost and is therefore discussed in detail in the present paper. The focus lies on the demonstration of the capabilities of this novel approach to adaptive refinement. Hence, the modeling of the FSW process is kept as simple as possible. To assess the results several numerical examples are presented.

The outline of the present paper is as follows: In Section 2 the governing equations for the mechanical problem are introduced and the visco-plastic material model is introduced in Section 2.2. Section 2.3 introduces the weak form and element stiffness matrices for the $\mathbf{u}/p$ formulation (Section 2.3.1) and the $\mathbf{u}/p/\epsilon$ formulation (Section 2.3.2), respectively. In Section 3 the thermal problem is introduced. In Section 4 the AFR approach is introduced, including the error estimators (Section 4.1), the adaptive approach (Section 4.3) as well as the solution strategy (Section 4.4). The results are presented and assessed in Section 5 and conclusions are drawn in Section 6.

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

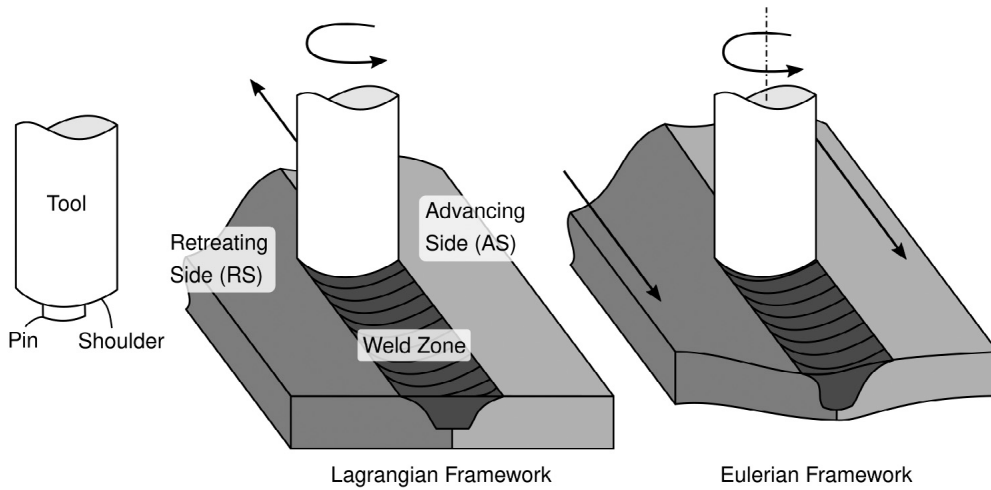


Fig. 1. Friction Stir Welding Process.

2. Mechanical analysis

In Friction Stir Welding a pin-tool is rotating around its axis while moving forward along the welding line. Thus, assuming a Lagrangian framework, it is necessary to perform continuous remeshing to adjust the mesh according to the current tool position. Alternatively, it is possible to consider a relative movement of the metal sheets towards the actual pin which is kept rotating around a fixed axis. In this case, the problem is solved in an Eulerian framework [29] (see Fig. 1).

The advantage of the Eulerian framework is the minimization of remeshing, particularly when a cylindrical tool is used. Bussetta et al. [10] showed that both frameworks yield equivalent results, though the Eulerian framework is generally preferred because it is computationally more efficient. The main consequence of adopting an Eulerian framework is related to the constitutive model used to characterize thermo-mechanical material behavior. The classical thermo-elasto-visco-plastic model, used within the Lagrangian framework, is typically replaced by a thermo-visco-plastic flow model, as introduced in Section 2.2.

2.1. Continuum problem

Let Ω denote the open and bounded domain in $\mathbb{R}^n$ occupied by the body defined in a space of n dimensions. The boundary Γ of Ω is split into Γ_u , where the prescribed velocities $\bar{\mathbf{u}}$ are specified (Dirichlet boundary conditions), and Γ_t , where the prescribed tractions $\bar{\mathbf{t}}$ are applied (Neumann boundary conditions), such that $\Gamma_u \cup \Gamma_t = \Gamma$ and $\Gamma_u \cap \Gamma_t = \emptyset$.

The mechanical model is described by the following set of equations

$$\nabla \cdot \mathbf{s} + \nabla p + \mathbf{b} = \mathbf{0} \quad (1)$$

$$\mathbf{e} - \nabla^s \mathbf{u} = \mathbf{0} \quad (2)$$

$$\mathbf{s} - 2\mu_{\text{eff}} \mathbf{e} = \mathbf{0} \quad (3)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (4)$$

where Eq. (1) is the balance of momentum equation, Eq. (2) the kinematic equation, Eq. (3) the constitutive equation, and, finally, Eq. (4) enforces the incompressibility constraint. In the above equations, $\mathbf{u}$ denotes the velocity field, being the driving variable of the problem, p and $\mathbf{s}(\mathbf{e})$ are the spherical (pressure) and the deviatoric parts of the stress tensor, μ_{eff} is the effective viscosity and $\mathbf{b}$ are the body forces. The strain rate tensor, $\mathbf{e} = \nabla^s \mathbf{u}$, is purely deviatoric due to the isochoric nature of this problem (incompressible flow). Observe that the convective term of the balance of momentum equation is neglected because in FSW the material flow is characterized by very low Reynolds numbers (the viscous forces being much larger than the inertial forces [12,29]).

2.2. Norton–Hoff material model

Visco-plastic behavior is widely used to describe the material flow for the simulation of several metal forming processes such as metal casting, extrusion or forging. Such laws neglect the elastic response and assume an incompressible material flow. Common visco-plastic material models are the Sheppard–Wright model [55], the (regularized) Bingham model [56], the Carreau model [57] and the Norton–Hoff model [58,59], which is a generalization of a non-Newtonian fluid, modeled as a power law. In the present work, the Norton–Hoff model is used because of the availability of temperature-dependent material parameters (see Fig. 19), obtained from an extensive experimental program.

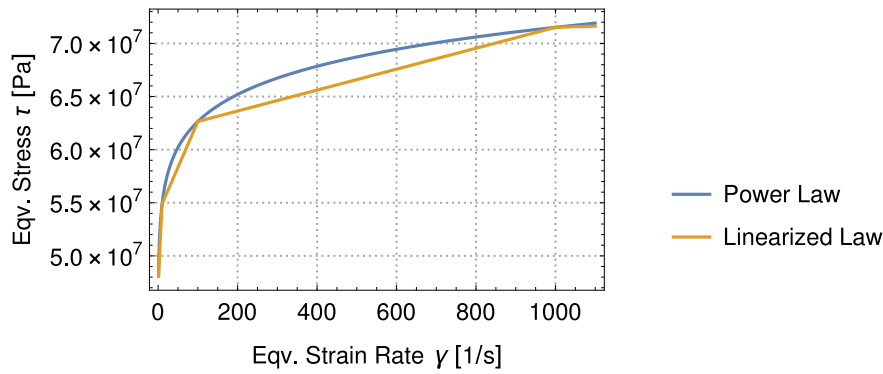


Fig. 2. Norton–Hoff power law and linearized law at a fixed temperature of $T = 300\text{ }^{\circ}\text{C}$.

The stresses $\mathbf{s}$ are obtained as

$$\mathbf{s} = 2\mu_{\text{eff}}\mathbf{e} \quad \text{with} \quad \mu_{\text{eff}}(\gamma) = \mu_0\gamma^{m-1}, \quad (5)$$

where the effective viscosity μ_{eff} is computed from the (temperature-dependent) consistency parameter μ_0 , the equivalent strain rate $\gamma = \sqrt{2}\|\mathbf{e}\|$ and the (temperature-dependent) flow index m . Note that for $m = 1$ Newtonian behavior is recovered, thus $\mu_{\text{eff}} = \mu_0$. In the present work, the flow index m ranges from 0.02 to 0.2, resulting in very non-linear behavior and leading to a *shear-thinning* effect at the boundary layer.

The mechanical dissipation D_m is computed as

$$D_m = \varphi \mathbf{s} : \mathbf{e}, \quad (6)$$

where φ is the fraction of the mechanical dissipation turned into heat.

Analogous to the equivalent strain rate, an equivalent stress τ can be computed as $\tau = \frac{\sqrt{2}}{2}\|\mathbf{s}\|$. Thus, the constitutive equation can also be written as $\tau = \mu_{\text{eff}}\gamma$ and the mechanical dissipation reduces to $D_m = \varphi\tau\gamma$.

Linearized law. Due to the highly non-linear nature of the constitutive equation for very low values of the flow index parameter m , a linearized version of the Norton–Hoff law is proposed. For each order of magnitude of the equivalent strain rate γ , a modified effective viscosity $\bar{\mu}_{\text{eff}}$ is computed by interpolating the values obtained at the two extremes $\mu_{\text{eff}1}$ and $\mu_{\text{eff}2}$.

First, the order of magnitude of the equivalent strain rate is determined to the smaller whole number ψ as

$$\psi = \lfloor \log_{10} \gamma \rfloor \quad \text{for} \quad \gamma > 1. \quad (7)$$

Then, two equivalent strain rates are computed as $\gamma_1 = 10^\psi$ and $\gamma_2 = 10^{\psi+1}$ and, according to Eq. (5), two effective viscosities $\mu_{\text{eff}1}$ and $\mu_{\text{eff}2}$ are computed. Finally, by performing a linear interpolation, the value of the modified effective viscosity $\bar{\mu}_{\text{eff}}$ is obtained. For values of $\gamma \leq 1$, the linearized effective viscosity is computed as $\bar{\mu}_{\text{eff}} = \mu_{\text{eff}}(\gamma = 1)$.

Fig. 2 shows the equivalent stress/strain rate curves of both the original and the linearized Norton–Hoff model for a constant temperature T .

2.3. Mixed finite element formulations

Two of the main challenges of simulating FSW are the isochoric material flow and the highly non-linear material behavior induced by the thermal softening. In the present work, mixed formulations are used to address both challenges. The mixed velocity/pressure ($\mathbf{u}/p$) formulation is typically used in CFD analysis to deal with *fully* incompressible fluid flows. This formulation allows for the use of arbitrary element types such as hexahedral, tetrahedral or prism elements as well as arbitrary orders of shape functions. Nevertheless, linear–linear shape functions are generally used to interpolate both the velocity and pressure fields [44]. Alternatively, the mixed velocity/pressure/deviatoric strain rate ($\mathbf{u}/p/\mathbf{e}$) formulation can be adopted. This three-field formulation has shown a superior stress/strain rate accuracy while also dealing with fully incompressible behavior [12]. However, the $\mathbf{u}/p/\mathbf{e}$ formulation is also computationally more expensive because it introduces the deviatoric strain rate tensor as a primary unknown, together with the velocity and the pressure fields.

The main objective of the present work is to enhance the efficiency of FSW simulations by using the $\mathbf{u}/p/\mathbf{e}$ formulation only in the regions of the computational domain where the enhanced stress/strain rate accuracy is required, while using the $\mathbf{u}/p$ formulation elsewhere.

In the following, the weak forms and element stiffness matrices of the two formulations are presented. The *Variational Multiscale Method* (VMS) [28] is used to stabilize both formulations, which is required because linear approximations are adopted for all unknowns.

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

2.3.1. Two-field Formulation – $\mathbf{u}/p$

The mixed $\mathbf{u}/p$ formulation is obtained by substituting the kinematic equation, Eq. (2), and the constitutive equation, Eq. (3), into the balance of momentum equation, Eq. (1). The resulting system of equations reads

$$\nabla \cdot (2\mu_{\text{eff}} \nabla^s \mathbf{u}) + \nabla p + \mathbf{b} = \mathbf{0} \quad (8)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (9)$$

Eq. (8) is the balance of momentum equation in mixed form, defined in terms of both the velocity field $\mathbf{u}$ and the pressure field p . Eq. (9) is the kinematic equation, used to enforce the incompressibility constraint.

Galerkin weak form. The corresponding weak form (after integration by parts) reads

$$\int_{\Omega} \nabla^s \delta \mathbf{u} \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u} dV + \int_{\Omega} (\nabla \cdot \delta \mathbf{u}) p dV + \int_{\Omega} \delta p (\nabla \cdot \mathbf{u}) dV = F(\delta \mathbf{u}) \quad (10)$$

where the work of the external loads is

$$F(\delta \mathbf{u}) = \int_{\Omega} \delta \mathbf{u} \cdot \mathbf{b} dV + \int_{\Gamma_t} \delta \mathbf{u} \cdot \bar{\mathbf{t}} dA, \quad (11)$$

and where $\delta \mathbf{u}$ and δp are the test functions. Note that the velocity test function vanishes on the boundary $\delta \mathbf{u} = 0$ on Γ_u . The problem is complemented by the Dirichlet and Neumann boundary conditions that are defined as

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{on} \quad \Gamma_u \quad (12)$$

$$\mathbf{s} \cdot \mathbf{n} = \bar{\mathbf{t}} \quad \text{on} \quad \Gamma_t, \quad (13)$$

where $\bar{\mathbf{u}}$ and $\bar{\mathbf{t}}$ are the prescribed velocities and tractions, respectively, and $\mathbf{n}$ is the unit outward normal to the surface Γ_t .

The discrete Galerkin Finite Element approximation can be obtained by discretizing the computational domain into Finite Elements. The solution of the continuum problem $\{\mathbf{u}, p\}$ is then approximated by $\{\mathbf{u}_h, p_h\}$. The discrete counterpart of the weak form therefore reads

$$\int_{\Omega^e} \nabla^s \delta \mathbf{u}_h \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u}_h dV + \int_{\Omega^e} (\nabla \cdot \delta \mathbf{u}_h) p_h dV + \int_{\Omega^e} \delta p_h (\nabla \cdot \mathbf{u}_h) dV = F(\delta \mathbf{u}_h) \quad (14)$$

Stabilization. If the same order of interpolation is used for both velocity and pressure fields, the mixed problem violates the LBB condition [60] and therefore lacks stability. To stabilize the formulation, the Variational Multiscale Method (VMS) is used. The stabilized $\mathbf{u}/p$ formulation reads

$$\int_{\Omega^e} \nabla^s \delta \mathbf{u}_h \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u}_h dV + \int_{\Omega^e} (\nabla \cdot \delta \mathbf{u}_h) p_h dV + \int_{\Omega^e} \delta p_h (\nabla \cdot \mathbf{u}_h) dV - \tau_u \int_{\Omega^e} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] dV = F(\delta \mathbf{u}_h), \quad (15)$$

where τ_u is a stabilization parameter [29] that is computed as

$$\tau_u = c_u \frac{h^2}{2\mu_{\text{eff}}}. \quad (16)$$

Remark 1. The operator $\mathbf{P}_h^\perp = \mathbf{I} - \mathbf{P}_h$ denotes the projection onto the space orthogonal to the Finite Element space, where $\mathbf{P}_h$ is the projection onto the appropriate Finite Element space. Typically the projection $\mathbf{P}_h$ is evaluated at the previous iteration and can therefore be moved to the right-hand side of the problem. A detailed derivation of the stabilized mixed $\mathbf{u}/p$ formulation for solid mechanics can be found in [29]. See also [61–63].

Element stiffness matrix. The element-wise approximations of the two-field problem are $\mathbf{u}_h = \mathbf{N}_u \mathbf{U}_h$ and $p_h = \mathbf{N}_p \mathbf{P}_h$ with $\mathbf{U}_h$ and $\mathbf{P}_h$ being the corresponding nodal values of the velocity and pressure fields, respectively. $\mathbf{N}_u$ and $\mathbf{N}_p$ are the matrices of the corresponding shape functions. The Galerkin method defines the test function by using the same shape functions used for the interpolation of the nodal variables, thus $\delta \mathbf{u}_h = \mathbf{N}_u \delta \mathbf{U}_h$ and $\delta p_h = \mathbf{N}_p \delta \mathbf{P}_h$. Thus, the corresponding elemental stiffness matrix reads

$$\mathbf{K} = \int_{\Omega^e} \begin{bmatrix} 2\mu_{\text{eff}} [\mathbf{B}^T \mathbf{B}] & [\mathbf{G} \mathbf{N}] \\ [\mathbf{N}^T \mathbf{G}^T] & -\tau_u [\mathbf{G}^T \mathbf{G}] \end{bmatrix} dV. \quad (17)$$

2.3.2. Three-field Formulation – $\mathbf{u}/p/e$

The mixed $\mathbf{u}/p/e$ formulation is obtained by introducing only the constitutive equation Eq. (3) into the balance of momentum equation Eq. (1) while preserving both the kinematic equation Eq. (2) and the incompressibility constraint Eq. (4).

$$\nabla \cdot (2\mu_{\text{eff}} \mathbf{e}) + \nabla p + \mathbf{b} = \mathbf{0} \quad (18)$$

$$\mathbf{e} - \nabla^s \mathbf{u} = \mathbf{0} \quad (19)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (20)$$

The mechanical problem is now formulated in terms of the deviatoric strain rate $\mathbf{e}$, the pressure p and the velocity $\mathbf{u}$ as primary unknowns.

Galerkin weak form. The corresponding weak form (after integration by parts) reads

$$\int_{\Omega} (\nabla \cdot \delta \mathbf{u}) \cdot p \, dV + \int_{\Omega} \nabla^s \delta \mathbf{u} \cdot 2\mu_{\text{eff}} \mathbf{e} \, dV + \int_{\Omega} \delta p (\nabla \cdot \mathbf{u}) \, dV + \int_{\Omega} \delta \mathbf{e} \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u} \, dV - \int_{\Omega} \delta \mathbf{e} \cdot 2\mu_{\text{eff}} \mathbf{e} \, dV = F(\delta \mathbf{u}), \quad (21)$$

where $\delta \mathbf{u}$, δp and $\delta \mathbf{e}$ are the test functions with $\delta \mathbf{u} = 0$ on Γ_u . Note that the third equation has been multiplied with $2\mu_{\text{eff}}$ to achieve symmetry. The problem is complemented by the same boundary conditions as for the $\mathbf{u}/p$ formulation.

In the corresponding discrete Galerkin Finite Element problem, the solution of the continuum problem $\{\mathbf{u}, p, \mathbf{e}\}$ is now approximated by $\{\mathbf{u}_h, p_h, \mathbf{e}_h\}$.

$$\int_{\Omega^e} (\nabla \cdot \delta \mathbf{u}_h) p_h \, dV + \int_{\Omega^e} \nabla^s \delta \mathbf{u}_h \cdot 2\mu_{\text{eff}} \mathbf{e}_h \, dV + \int_{\Omega^e} \delta p_h (\nabla \cdot \mathbf{u}_h) \, dV + \int_{\Omega^e} \delta \mathbf{e}_h \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u}_h \, dV - \int_{\Omega^e} \delta \mathbf{e}_h \cdot 2\mu_{\text{eff}} \mathbf{e}_h \, dV = F(\delta \mathbf{u}_h) \quad (22)$$

Stabilization. As for the $\mathbf{u}/p$ formulation, the VMS method is adopted to stabilize the formulation. The stabilized $\mathbf{u}/p/\mathbf{e}$ formulation reads

$$\begin{aligned} \int_{\Omega^e} (\nabla \cdot \delta \mathbf{u}_h) p_h \, dV + \int_{\Omega^e} \nabla^s \delta \mathbf{u}_h \cdot 2\mu_{\text{eff}} \mathbf{e}_h \, dV + \int_{\Omega^e} \delta p_h (\nabla \cdot \mathbf{u}_h) \, dV + \int_{\Omega^e} \delta \mathbf{e}_h \cdot 2\mu_{\text{eff}} \nabla^s \mathbf{u}_h \, dV - \int_{\Omega^e} \delta \mathbf{e}_h \cdot 2\mu_{\text{eff}} \mathbf{e}_h \, dV \\ + \tau_e \int_{\Omega^e} \nabla^s \delta \mathbf{u}_h \cdot \mathbf{P}_h^\perp [2\mu_{\text{eff}} \nabla^s \mathbf{u}_h] \, dV - \tau_u \int_{\Omega^e} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV = F(\delta \mathbf{u}_h), \end{aligned} \quad (23)$$

where the stabilization parameters [46] are computed as

$$\tau_u = c_u \frac{h^2}{2\mu_{\text{eff}}} \quad (24)$$

$$\tau_e = c_e. \quad (25)$$

Remark 2. Analogous to the stabilized $\mathbf{u}/p$ formulation, the stabilization terms for the $\mathbf{u}/p/\mathbf{e}$ also includes projections onto the space orthogonal to the Finite Element space, which also can be moved to the right-hand side of the equation, when evaluated at the previous iteration. For detailed derivation of the stabilized mixed $\mathbf{u}/p/\mathbf{e}$ formulation see [12,46,50].

Element stiffness matrix. The element-wise approximations of the three solution fields are $\mathbf{u}_h = \mathbf{N}_u \mathbf{U}_h$, $p_h = \mathbf{N}_p \mathbf{P}_h$ and $\mathbf{e}_h = \mathbf{N}_e \mathbf{E}_h$ with $\mathbf{U}_h$, $\mathbf{P}_h$ and $\mathbf{E}_h$ being the corresponding nodal values and $\mathbf{N}_u$, $\mathbf{N}_p$ and $\mathbf{N}_e$ being the corresponding matrices of shape functions. The test functions are defined as $\delta \mathbf{u}_h = \mathbf{N}_u \delta \mathbf{U}_h$, $\delta p_h = \mathbf{N}_p \delta \mathbf{P}_h$ and $\delta \mathbf{e}_h = \mathbf{N}_e \delta \mathbf{E}_h$. Thus, the corresponding elemental stiffness matrix reads

$$\mathbf{K} = \int_{\Omega^e} \begin{bmatrix} \tau_e 2\mu_{\text{eff}} [\mathbf{B}^\top \mathbf{B}] & [\mathbf{G}\mathbf{N}] & 2\mu_{\text{eff}} [\mathbf{B}^\top \mathbf{N}] \\ [\mathbf{N}^\top \mathbf{G}^\top] & -\tau_u [\mathbf{G}^\top \mathbf{G}] & [\mathbf{0}] \\ 2\mu_{\text{eff}} [\mathbf{N}^\top \mathbf{B}] & [\mathbf{0}] & -2\mu_{\text{eff}} [\mathbf{N}^\top \mathbf{N}] \end{bmatrix} dV. \quad (26)$$

3. Thermal analysis

3.1. Continuum problem

The governing equation of the thermal problem reads

$$\rho_0 c \left(\frac{1}{\alpha} \frac{dT}{dt} + \mathbf{u} \cdot \nabla T \right) + \nabla \cdot \mathbf{q} = D_m \quad \text{with} \quad \mathbf{q} = -k \nabla T, \quad (27)$$

where T is the temperature field (the driving variable of the problem), $\mathbf{u}$ the convective velocity, ρ_0 the material density, c is the (temperature-dependent) specific heat capacity and $\mathbf{q}$ is the heat flux, where k is the (temperature-dependent) thermal conductivity. D_m denotes the plastic dissipation (see Section 2.2), acting as source term.

The parameter α is introduced to artificially accelerate the transient phase by increasing the thermal diffusion. Thus, the steady state condition can be achieved faster. This method was first introduced by Dialami et al. [64] for the numerical simulation of FSW processes.

3.2. Finite element formulation

The corresponding weak form of the thermal problem reads

$$\int_{\Omega} \delta T \rho_0 c \left(\frac{1}{\alpha} \frac{dT}{dt} + \mathbf{u} \cdot \nabla T \right) dV + \int_{\Omega} \nabla \delta T \cdot k \nabla T \, dV = \int_{\Omega} \delta T \cdot D_m \, dV - \int_{\Gamma_q} \delta T \cdot \bar{q} \, dA, \quad (28)$$

where δT is the test function of the temperature field.

The problem is complemented by the boundary conditions. The boundary Γ of Ω is split into Γ_T , where the prescribed temperatures $\bar{T}$ are specified (Dirichlet boundary conditions), and Γ_q , where the prescribed heat fluxes $\bar{q}$ are applied (Neumann boundary conditions), such that $\Gamma_T \cup \Gamma_q = \Gamma$ and $\Gamma_T \cap \Gamma_q = \emptyset$. The boundary conditions are defined as

$$T = \bar{T} \quad \text{on} \quad \Gamma_T \quad (29)$$

$$\mathbf{q} \cdot \mathbf{n} = \bar{q} \quad \text{on} \quad \Gamma_q, \quad (30)$$

where $\mathbf{n}$ is the unit outward normal to the surface Γ_q . On surfaces which are in contact with the back plate, the heat is dissipated by conduction and can be expressed by Newton's law as

$$q_{\text{cond}} = h_{\text{cond}}(T - T_{\text{backplate}}), \quad (31)$$

where h_{cond} is the heat transfer coefficient by conduction and $T_{\text{backplate}}$ the environment temperature. Heat fluxes to the environment by convection and radiation are not considered. The initial temperature for the transient thermal problem is $T(t=0) = T_0$.

The discrete Galerkin Finite Element approximation can be obtained by discretizing the computational domain into Finite Elements. The solution of the continuum problem T is now approximated by T_h . Thus, the discrete counterpart of the weak form reads

$$\int_{\Omega^e} \delta T_h \rho_0 c \left(\frac{1}{\alpha} \frac{dT_h}{dt} + \mathbf{u}_h \cdot \nabla T_h \right) dV + \int_{\Omega^e} \nabla \delta T_h \cdot k \nabla T_h dV = \int_{\Omega^e} \delta T_h D_m dV - \int_{\Gamma_q^e} \delta T_h \bar{q} dA. \quad (32)$$

Remark 3. The Eulerian formulation, adopted for the mechanical problem, introduces convection terms in the governing equations of the thermal problem. In convection dominated problem this may cause stability problems. In this work, Galerkin Least-Squares (GLS) stabilization is used to stabilize the thermal problem [65].

Remark 4. The coupled thermo-mechanical problem is solved through a staggered scheme derived from a fractional step method [1]. Thus, the mechanical problem is solved first, based on the current temperature field. Next, the thermal field is computed, based on the updated velocity field. Note that the mechanical problem requires temperature-dependent material properties (i.e., the viscosity and flow index), while the thermal problem requires the velocity field $\mathbf{u}$ to compute the convection term and the mechanical dissipation D_m as heat source.

4. Adaptive formulation refinement

In order to define a most suitable and computationally efficient framework for an FSW simulation, the adaptive use of both the two- and three-field formulations is proposed in this work. The AFR method switches between the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations, driven by a refinement criterion based on a posteriori error estimation.

When compared to established refinement approaches (i.e., adaptive h - and p -refinement), the AFR method shows some advantages which are beneficial for simulating FSW problems. By using a stabilized mixed formulation, linear interpolation can be adopted for all variable fields. Moreover, the same mesh is used during the whole analysis, thus no remeshing is required. Finally, the treatment of the interface between refined and unrefined subdomains is straightforward since no hanging nodes need to be considered. The AFR approach can be broken down into four parts:

1. Definition of an a posteriori error estimator, to drive the adaptive method
2. Strain rate and stress consistency
3. Definition of the adaptivity approach
4. Integration into a High Performance Computing (HPC) framework

4.1. Refinement criteria and error estimation

For some applications it is possible to know a priori where a high accuracy is required. Thus, a finer mesh, a different formulation or higher order shape functions, can be set at the beginning of the simulation. However, this is typically not known a priori and adaptive approaches are used instead. As a consequence, a posteriori error estimators and refinement criteria are used to determine where the highest errors occur and refinement is required. For the AFR approach, this means determining where to use the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations, respectively.

In the present work, the well known Zienkiewicz–Zhu (ZZ) error estimator [52] is used. The local (element-wise) error e_{ZZ} is estimated by comparing two strain rates ($\mathbf{e}_p$ and $\mathbf{e}_c$) at the center of the element. The first strain rate $\mathbf{e}_p$ is obtained by computing the symmetric gradient of the velocity at the integration points, projecting them to the nodes using a lumped mass matrix projection and interpolating the results to the center of the element. The second strain rate $\mathbf{e}_c$ is computed as the symmetric gradient of the velocity field directly at the center of the element without projecting it to the nodes first. Subsequently, the difference is computed as $\mathbf{e}_{\text{diff}} = \mathbf{e}_p - \mathbf{e}_c$ and the norm of the strain rate difference is integrated over the element according to Eq. (33).

$$e_{ZZ} = \int_{\Omega^e} \mathbf{e}_{\text{diff}} : \mathbf{e}_{\text{diff}} d\Omega^e \quad (33)$$

For Finite Element approaches with C^0 continuity, this error estimator approximately quantifies the difference between the strain rate field obtained by computing the symmetric gradient of the velocity field in the $\mathbf{u}/p$ formulation and the strain rate field that would be obtained by using the $\mathbf{u}/p/e$ approach. Hence, it is a good indicator to switch from $\mathbf{u}/p$ to $\mathbf{u}/p/e$.

4.2. Strain rate and stress consistency

This subsection aims to explain how strain rates and stresses are obtained in the different mixed formulations. This is important to later ensure a consistency of those fields at the interface between unrefined ($\mathbf{u}/p$) and refined ($\mathbf{u}/p/e$) subdomains.

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

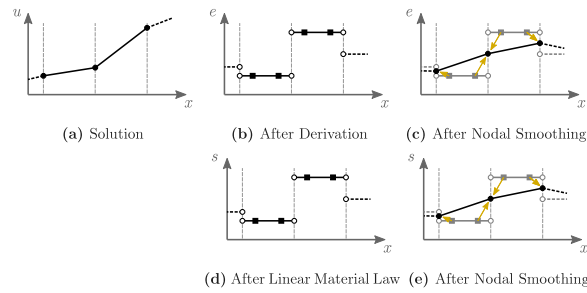


Fig. 3. Strain rate and stress handling in two-field u/p formulation.

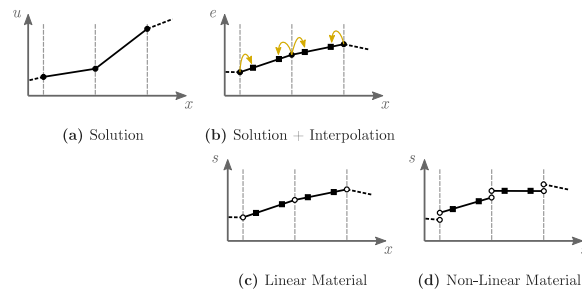


Fig. 4. Strain rate and stress handling in three-field $u/p/e$ formulation.

u/p formulation. In the two-field formulation, the velocity field $\mathbf{u}$ is computed at the nodes (●, Fig. 3(a)). Assuming linear elements, the resulting field is therefore C^0 -continuous. To obtain the strain rate $\mathbf{e}$, the first derivative is computed at the integration points (■) of the elements (Fig. 3(b)). Due to the derivation, the strain rate field becomes C^{-1} -continuous or *discontinuous*. This is especially obvious for linear simplex elements since the strain rate is constant within each element. Finally the stresses are obtained (Fig. 3(d)) by evaluating the material law. Since the strain rates are computed and stored at the integration points, the material law is also evaluated at the integration points. Therefore, both the stresses and the mechanical dissipation are computed at the same location. To postprocess the strain rate, stress or dissipation, the fields are often projected onto the nodes (Figs. 3(c) and 3(e)) in order to show continuous fields.

$u/p/e$ formulation. Unlike in the two-field formulation, the strain rate is now considered an additional primary unknown and is therefore defined at the nodes. (Fig. 4(b)). This given, two different approaches could be followed.

The first approach would be to also compute stresses and the dissipation at the nodes. However, this would require to evaluate the material law at the nodes, which is possible but uncommon. It would also create an inconsistency between the u/p and $u/p/e$ subdomains.

To keep consistency between the u/p and $u/p/e$ subdomains, the same approach as for the two-field formulation is followed. The strain rate field is available at the nodes (●), hence it is C^0 -continuous. Afterwards the field is interpolated to the integration points (■, Fig. 4(b)) and thus it is still continuous. Considering a linear material law, also the stresses are continuous (Fig. 4(c)). Only when considering a non-linear material law, a discontinuous stress field can be observed (Fig. 4(d)). Finally, the mechanical dissipation, which is computed from the stresses and the strain rate, can also be easily computed at the integration points.

4.3. Adaptive approach

At the beginning of each time step, the evaluation of the refinement criterion (based on the chosen error estimator) yields the definition of the u/p and $u/p/e$ subdomains. The element stiffness matrices are then computed according to which subdomain the current element belongs to. Since the u/p formulation is a subset of the $u/p/e$ formulation, it is convenient to consider the three-field formulation as the default approach. Thus, in the u/p subdomain the degrees of freedom associated with the strain rate are deactivated.

At the interface between the u/p and $u/p/e$ elements all three-field variables ($\mathbf{u}$, p and $\mathbf{e}$) are preserved. A schematic representation of the interface between two-field and three-field elements can be seen in Fig. 5. Note that all nodes at the interface are $u/p/e$ nodes.

Thanks to the unified way of treating the strain rate, the *mixed* strain rate field can be obtained straightforwardly, as seen in Fig. 6. In the u/p subdomain the strain rate is obtained by computing the symmetric gradient of the velocity at the integration points, whereas in the $u/p/e$ subdomain, the strain rate is directly obtained as a solution at the nodes and afterwards interpolated to the

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

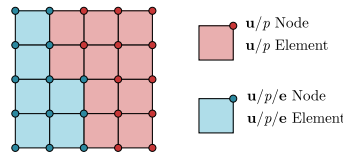


Fig. 5. Interface of $u/p/e$ and u/p elements and nodes.

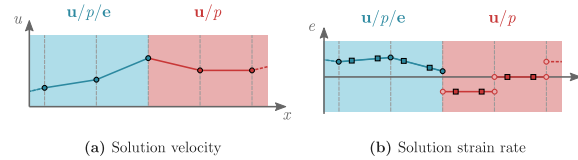


Fig. 6. Solution fields at the formulation interface.

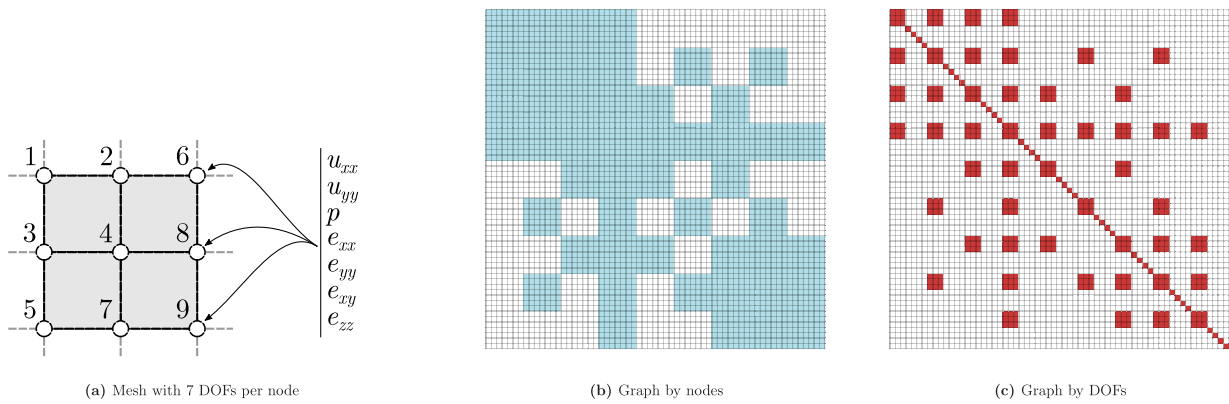


Fig. 7. Visualization of non-zero matrix entries of an AFR problem.

integration points. Therefore, the subsequent computation of the stresses and the dissipation can be carried out at the integration points.

Remark 5. Special attention must be paid to the computation of the strain rate subscales when using orthogonal sub-grid stabilization (OSS, a specific type of VMS stabilization [29,61]). The computation of the projection onto the space of subscales $\tilde{P}$ involves the nodal projection of the subscales of the affected variables. This poses no problem for the velocity and pressure subscales since both fields are defined on the whole domain. The strain rate subscales, however, only exist in the $u/p/e$ subdomain. Consequently, the nodes at the interface between the u/p and $u/p/e$ subdomains lack a contribution from adjacent u/p elements, when the nodal projection of the strain rate subscales is performed. To solve this problem the nodal strain rate in the u/p subdomain is approximated by a nodal projection of the symmetric gradient of the velocity.

4.4. Solver strategies

The purpose of the deactivation of the degrees of freedom (DOFs) associated to the strain rate in the u/p subdomain is to create a problem which is smaller (less required memory) and faster to solve (less required CPU-time). However, to achieve this, some aspects of the deactivation process requires further discussion.

In most Finite Element codes, the system matrices are usually stored in a compressed format like the *Compressed Sparse Row* (CSR) format. The graph of the matrix is used to locate an entry of the matrix. Note that the graph only accounts for non-zero entries of the matrix.

In the following, the small mesh shown in Fig. 7(a) is analyzed. In most Finite Element applications all nodes have the same number of DOFs and almost all assembled values are non-zeros. In such cases it is convenient to build the graph by nodes. This allows all DOFs associated with a particular node to be assembled in a single operation. The resulting pattern of non-zeros is shown in Fig. 7(b).

In the context of the AFR method, the mesh shown in Fig. 7(a) could correspond to a u/p subdomain. Hence, only the velocity and pressure DOFs are of interest. However, if no further action is taken, the graph and pattern of non-zeros would remain the same as the one corresponding to the default $u/p/e$ formulation. This causes two problems. First, a large number of zeros are assembled

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

```
-ksp_type bcgs
-ksp_rtol 1e-8
-ksp_max_it 300
-ksp_diagonal_scale
```

Fig. 8. PETSc solver settings. Note that the maximum number of solver iterations specified (300) is never reached. The default preconditioning in PETSc is ILU(0).

to the system, which is increasing the build time unnecessarily. Second, although the assembled zeros do not contribute to the solution, the solver needs to consider these entries when solving the system because – according to the graph – they are considered non-zeros. This increases the solve time.

The solution to this problem is to build the graph by DOFs. This allows to only consider the $\mathbf{u}/p$ DOFs as well as a unitary value in the positions of the main diagonal which correspond to strain rate components as part of the graph and therefore as non-zero entries. Hence, the latter do not contribute noticeably to the solving time since they are the only non-zero entries in their respective rows and columns. This results in faster build and solve times and avoids any renumbering of the DOFs since all matrices remain the same size.

The resulting pattern of non-zeros is shown in Fig. 7(c).

Remark 6. The activation and deactivation of DOFs is handled at runtime. For each time step the refinement criterion is evaluated and the elements (and respective nodes) are determined to be either $\mathbf{u}/p/e$ or $\mathbf{u}/p$. The DOF graph is then modified correspondingly.

Remark 7. The authors of the present paper use the PETSc library [66] to build and solve the system. In PETSc, the expression ‘BAIJ-matrix’ (B stands for ‘blocked’) refers to matrices that support multiple DOFs being assembled as one block in a single operation (i.e., the graph is built by nodes). In the example shown in Fig. 7(a), the block size is 7 (the number of DOFs per node), and must be the same for all nodes. To be able to vary the number of DOFs per node, the graph must be built by DOFs. Therefore, in PETSc, an ‘AIJ-matrix’ must be used instead.

5. Numerical examples

In this section the proposed AFR method is assessed by solving several numerical examples. Four different examples are shown, each giving a different perspective on the performance of the presented method.

Example 1 The build and solve times for different matrix types and assembly strategies are studied. This example assesses the results obtained with the PETSc solver library.

Example 2 A simple 2D flow over a cylinder case is presented. This example shows the dissipation field for the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations, as well as the solution using AFR. The solution accuracy in different parts of the domain is assessed with regard to the different formulations.

Example 3 A global error analysis is performed in order to study the convergence rates upon mesh refinement and to prove that the AFR method converges to the analytical solution. Furthermore, the influence of the arrangement of $\mathbf{u}/p$ and $\mathbf{u}/p/e$ subdomains is studied.

Example 4 Finally, an FSW application is presented. The example is a 3D, thermo-mechanically coupled case with highly non-linear material behavior. This final example aims to demonstrate the use of the AFR method for industrial applications.

In all examples, the stabilization coefficients (see Eqs. (17) and (26)) are $c_u = 1.0$ and $c_e = 0.1$ for both the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations.

5.1. Build and solve times assessment

To measure the build and solve times, a 3D case consisting of 70 490 nodes and 66 470 hexahedral elements, is studied for different formulations and matrix types. In the $\mathbf{u}/p$ cases, this resulted in building and solving a system with 281 960 unknowns (number of nodes times 4 DOF: u_x, u_y, u_z, p) and in $\mathbf{u}/p/e$ and AFR cases, a system with 704 900 unknowns (number of nodes times 10 DOF: $u_x, u_y, u_z, p, e_{xx}, e_{yy}, e_{zz}, e_{xy}, e_{xz}, e_{yz}$). A single iteration is carried out in parallel (6 Threads) on an Intel Core i7-10750H CPU. For this analysis, the PETSc implementation of the stabilized version of the biconjugate gradient method (BCGS) is used, with the settings shown in Fig. 8.

The build and solve times for two different matrix types (AIJ and BAIJ) and assembly strategies (node graph and DOF graph), as introduced in Section 4.4, are shown in Table 1. In the first section (labeled *genuine formulation*, Nos. 1 and 2) no DOFs are activated nor deactivated. Only the straightforward implementations of the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations are used. In both cases the

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

Table 1
Build and solve times for different methods and matrix types.

No.	Method, Matrix type, Graph type		Unknowns	Build time	Solve time
1	genuine formulation, BAIJ, node graph	$\mathbf{u}/p$	281 960	0.85 s	7.75 s
		$\mathbf{u}/p/e$	704 900	3.15 s	62.97 s
2	genuine formulation, AIJ, node graph	$\mathbf{u}/p$	281 960	1.05 s	10.02 s
		$\mathbf{u}/p/e$	704 900	4.98 s	71.65 s
3	AFR method, AIJ, node graph	$\mathbf{u}/p$	704 900	3.31 s	54.58 s
		$\mathbf{u}/p/e$	704 900	4.72 s	73.21 s
4	AFR method, AIJ, DOF graph	$\mathbf{u}/p$	704 900	→ 1.75 s	→ 12.30 s
		$\mathbf{u}/p/e$	704 900	4.72 s	71.66 s

Fig. 9. Flow over a cylinder: analysis of dissipation.

graph was built by nodes. The build and solve times of the genuine formulations indicate the range in which the times of the AFR method should fall. Considering the solve times first, the lower bound is set by the genuine $\mathbf{u}/p$ formulation with 7.75 s and 10.02 s for the BAIJ and AIJ matrices, respectively. The upper bound is set by the genuine $\mathbf{u}/p/e$ formulation with 62.97 s and 71.65 s, respectively. As expected, it takes more time to solve the $\mathbf{u}/p/e$ problem than the $\mathbf{u}/p$ problem and that both building and solving are faster if BAIJ matrices are used.

The second section (Nos. 3 and 4) shows the results obtained by using the AFR method. This means that the system is always of the size of the corresponding $\mathbf{u}/p/e$ problem and in the $\mathbf{u}/p$ cases, the DOFs associated with the strain rate were not assembled. The goal of using the AFR method is to achieve build and solve times that are as close as possible to those of the genuine $\mathbf{u}/p$ formulation, despite dealing with a system of the size of the $\mathbf{u}/p/e$ problem.

This is possible when the AFR method uses an AIJ matrix in order to take advantage of the DOF graph (No. 4). The performance of the AFR method can be assessed by comparing the $\mathbf{u}/p$ (AIJ) times of the AFR method (No. 4, 12.30 s) to the $\mathbf{u}/p$ (AIJ) and $\mathbf{u}/p/e$ (AIJ) times of the genuine formulations (10.02 s and 71.65 s). The AFR method is only marginally slower than the genuine $\mathbf{u}/p$ formulation, despite solving a system with 2.5 times as many unknowns.

The necessity of using a DOF graph instead of a node graph becomes obvious when comparing the AFR times for different graph types (Nos. 3 and 4): 12.30 s vs. 54.58 s.

Note that the build and solve times of the $\mathbf{u}/p/e$ problem do not vary significantly between Nos. 2 to 4. This shows that the implementation of the AFR method does not introduce any notable computational overhead when dealing with a *full* problem where all possible DOFs are active.

Using a graph based on the active DOFs results in a faster solution times compared to the original node-based graph. The build times follow a similar pattern, though for big problems the solving time dominates the overall required CPU-time.

5.2. Local error analysis

The second example assesses the local solution of a 2D *flow-over-a-cylinder* case. The geometry and boundary conditions are shown in Fig. 9. The square domain is of size 10 m × 10 m and the cylinder located in the center of the domain has a diameter of 2 m. The inflow velocity is set to $u_x = 0.01 \text{ ms}^{-1}$. The top and bottom boundaries as well as the boundary of the cylinder are modeled as *slip wall* conditions, thus the tangential velocity is free and the normal velocity is set to $u_n = 0$. This example is purely mechanical (temperature coupling is not accounted for) and adopts linear material behavior (Newtonian law). The material viscosity is set to $\mu_0 = 1 \cdot 10^8 \text{ Pa s}$ and the flow index to $m = 1$.

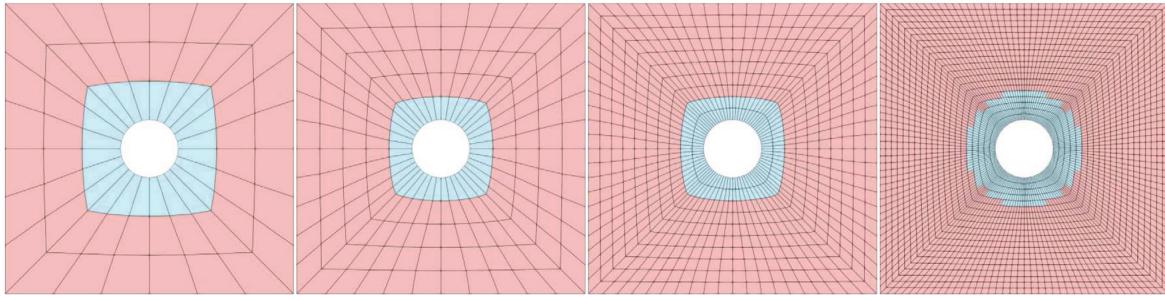


Fig. 10. Differently fine meshes: 96, 240, 880 and 3360 nodes.

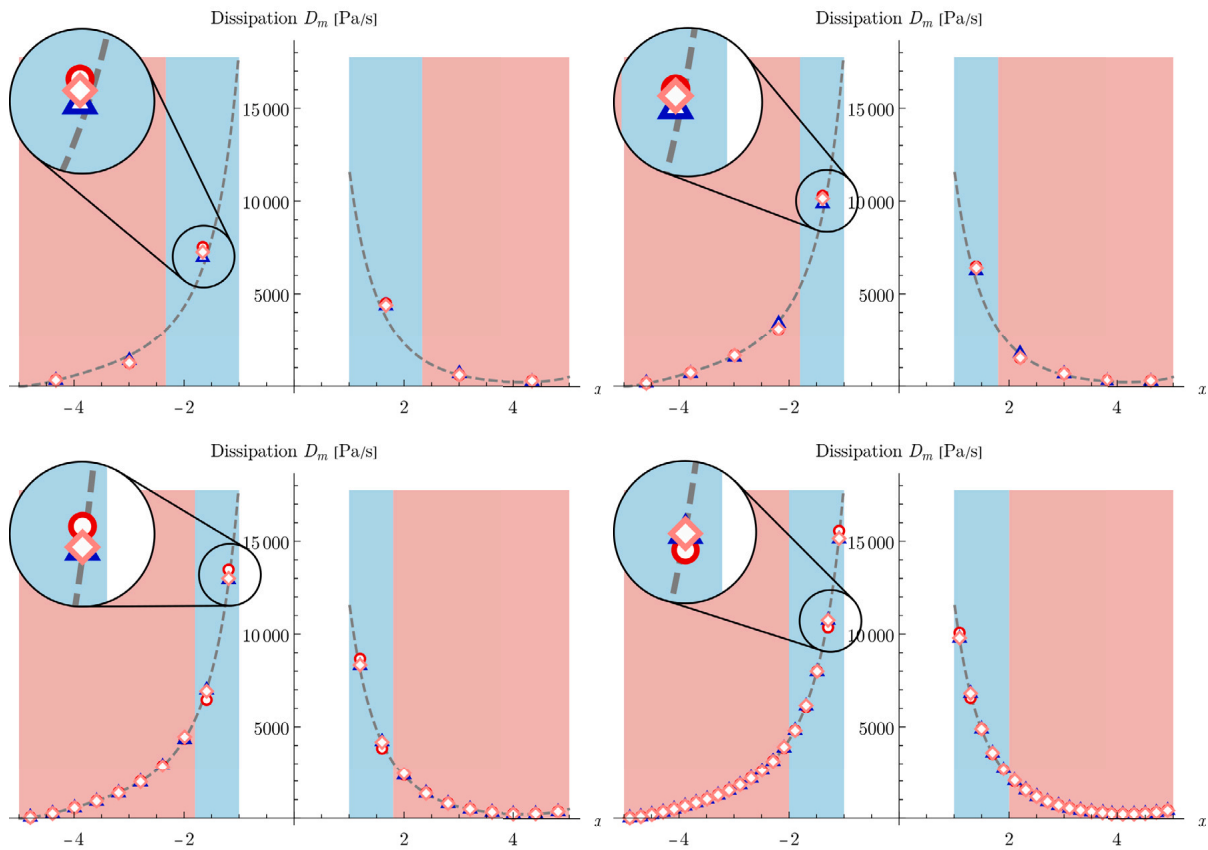


Fig. 11. Dissipation over line $A - A'$ for different mesh resolutions: reference solution indicated by dashed line, $\circ$ u/p , $\triangle$ $u/p/e$ and $\diamond$ AFR.

Fig. 9 shows a plot of the dissipation D_m on a very fine (120 800 nodes) reference mesh. Fig. 11 shows line plots ($A - A'$) of the dissipation for four gradually finer meshes (see Fig. 10). The dashed line indicates the reference solution as seen in Fig. 9. The different markers indicate the solution of the u/p formulation ($\circ$), the $u/p/e$ formulation ($\triangle$) and the AFR method ($\diamond$). In the case of the AFR method, the different subdomains are indicated by the background color: light red for the u/p subdomain and light blue for the $u/p/e$ subdomain. These subdomains were created a priori.

Far from the cylinder axis, where smaller dissipation values are expected, all three methods give very similar results, matching the reference solution. However, the dissipation field close to the cylinder shows a larger deviation from the reference solution. This deviation is always higher for the u/p formulation and is always lower for the $u/p/e$ formulation.

For all four meshes, the solution of the AFR strategy is located between the solution of the u/p and $u/p/e$ formulations. For the two coarser meshes, where only a single element is part of the $u/p/e$ subdomain, the solutions of the AFR strategy is closer to the u/p solution. Whereas, if more than one element belongs the $u/p/e$ subdomain, the AFR solution is almost identical to the $u/p/e$ solution.

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

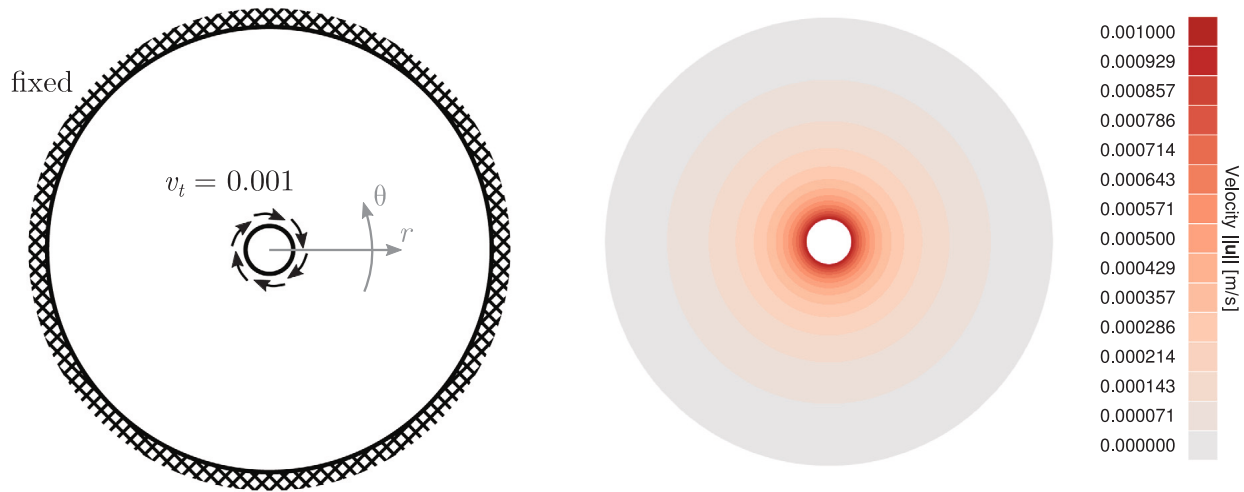


Fig. 12. Flow around cylinder: analytical solution of velocity norm $\|\mathbf{u}\|$.

The outcome is the following: (1) in low dissipation areas, all three formulations yield similar results, (2) the AFR solution seems to be bounded by the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ solutions and (3) if the $\mathbf{u}/p/e$ subdomain is sufficiently large, the AFR solution tends towards the $\mathbf{u}/p/e$ solution.

5.3. Global error analysis

In this example a convergence study upon mesh refinement is carried out. In the present study, the convergence rates of the velocity $\mathbf{u}$ and the energy e are examined. The norms of the errors [67] are obtained as follows

$$\|\mathbf{u} - \mathbf{u}_h\| = \left[\int_{\Omega} (\mathbf{u} - \mathbf{u}_h) : (\mathbf{u} - \mathbf{u}_h) d\Omega \right]^{1/2} \quad (34)$$

$$\|e - e_h\| = \left[\int_{\Omega} (\mathbf{s} - \mathbf{s}_h) : (\mathbf{e} - \mathbf{e}_h) d\Omega \right]^{1/2}, \quad (35)$$

where $\mathbf{u}$ is the velocity and $\mathbf{s}$ and $\mathbf{e}$ are the deviatoric stresses and strains in Voigt notation, respectively. The index $\square_h$ indicates the Finite Element approximation.

The geometry and boundary conditions are presented in Fig. 12. The circular domain has an outer diameter of 20 m and an inner diameter of 2 m. Homogeneous Dirichlet boundary conditions are applied to the outer perimeter of the circular domain. The inner perimeter is subjected to a prescribed velocity in tangential direction of $u_{\theta} = 0.001$ m/s. The resulting velocity field can be observed in Fig. 12. The example is purely mechanical (the temperature coupling was not modeled) and linear material behavior is adopted (Newtonian law). The material viscosity is set to $\mu_0 = 1$ Pa s and the flow index to $m = 1$.

For the given geometry and boundary conditions, the analytical solution is used to compute a global error in terms of the velocity field, its spacial derivatives and the pressure field:

$$u_r = 0 \quad u_{\theta} = -\frac{100 + r^2}{99\,000\,r} \quad (36)$$

$$u_{r,r} = u_{r,\theta} = u_{\theta,\theta} = 0 \quad u_{\theta,r} = -\frac{100 + r^2}{99\,000\,r^2} \quad (37)$$

$$p = \text{const.} \quad (38)$$

The AFR strategy relies on optimizing the use of the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ subdomains according to an error criterion. For the given example, the $\mathbf{u}/p/e$ formulation should be used close to the circular hole in the middle of the domain. The convergence study assesses the convergence rates for three differently subdivided domains and compares the results of the AFR method to the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations.

Fig. 13(a) shows a reasonable subdivision scheme, where the $\mathbf{u}/p/e$ subdomain is close to the center of the whole domain, whereas Figs. 13(b) and 13(c) depict two unreasonable subdivision schemes.

To assess the convergence rates, the global error is plotted over the mesh size h and the total number of degrees of freedom. Figs. 14 and 15 show the velocity error $\|\mathbf{u} - \mathbf{u}_h\|$ for different subdivision schemes, while Figs. 16 and 17 show the energy norm error $\|e - e_h\|$. All plots show the $\mathbf{u}/p$ (■) and $\mathbf{u}/p/e$ (■) results as reference.

It can be clearly seen from Figs. 14 and 15, that the velocity error $\|\mathbf{u} - \mathbf{u}_h\|$ decreases with the same convergence rate for all formulations. Though, for a given mesh size h , the $\mathbf{u}/p/e$ solution (■), as well as the AFR solution with the reasonable subdivision

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

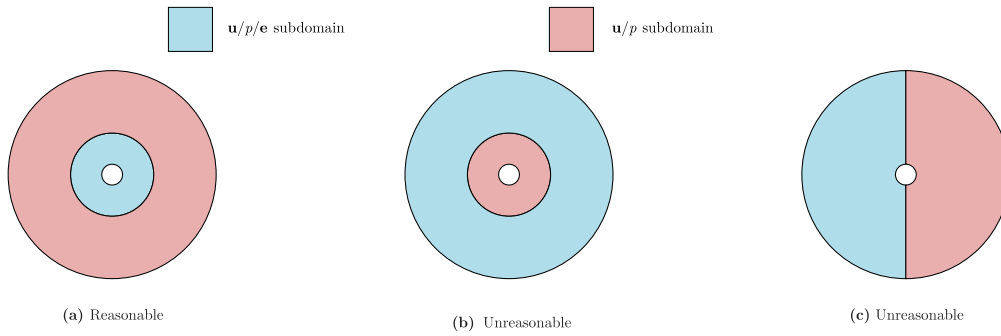


Fig. 13. Different strategies to divide the domain in $\mathbf{u}/p$ and $\mathbf{u}/p/e$ subdomains.

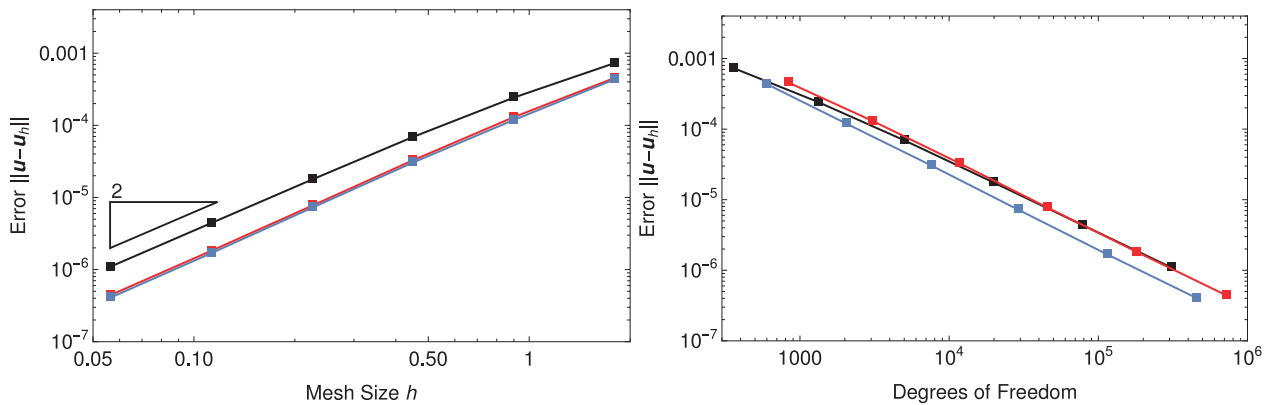


Fig. 14. Velocity errors $\|\mathbf{u} - \mathbf{u}_h\|$ for different subdivision schemes: $\mathbf{u}/p/e$, $\mathbf{u}/p$ and AFR (reasonable, Fig. 13(a)).

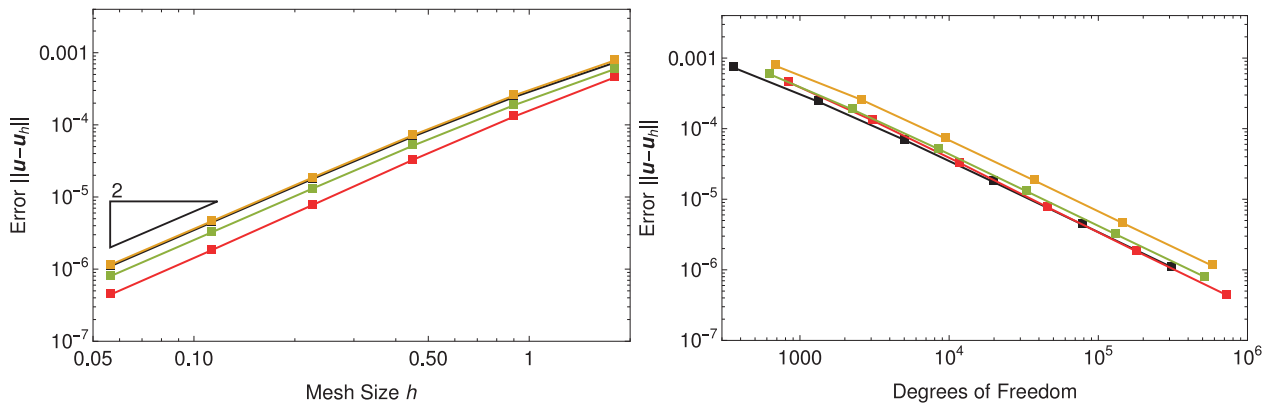


Fig. 15. Velocity errors $\|\mathbf{u} - \mathbf{u}_h\|$ for different subdivision schemes: $\mathbf{u}/p/e$, $\mathbf{u}/p$, AFR (unreasonable, Fig. 13(b)) and AFR (unreasonable, Fig. 13(c)).

scheme (■), yield more accurate results. However, if the error is displayed over the number of DOFs, the AFR solution (■) always yields the highest accuracy for a given number of DOFs. As expected, the AFR solutions using the unreasonable subdomains (■, ■) perform worse than the $\mathbf{u}/p$ formulation (■): for the same accuracy they require a higher number of DOFs.

A different behavior can be observed, when the error is studied using the energy norm $\|e - e_h\|$, as shown in Figs. 16 and 17. It can be observed that the $\mathbf{u}/p/e$ formulation (■) yields both the highest convergence rate and the most accurate results, while the $\mathbf{u}/p$ formulation (■) yields both the lowest convergence rate and the worst accuracy. The results (accuracy and convergence rate) obtained by the AFR technique are bounded by the results of the $\mathbf{u}/p$ and $\mathbf{u}/p/e$ formulations. The two unreasonable subdivision schemes (■, ■) perform as bad as the $\mathbf{u}/p$ formulation (or slightly better) while the reasonable subdivision scheme (■) performs almost as good as the $\mathbf{u}/p/e$ formulation. Only for very small mesh sizes, the convergence rate of the reasonable subdivision scheme (■) shifts towards the $\mathbf{u}/p$ behavior. This is expected, as for very fine meshes, the error of the $\mathbf{u}/p$ subdomain begins to dominate

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

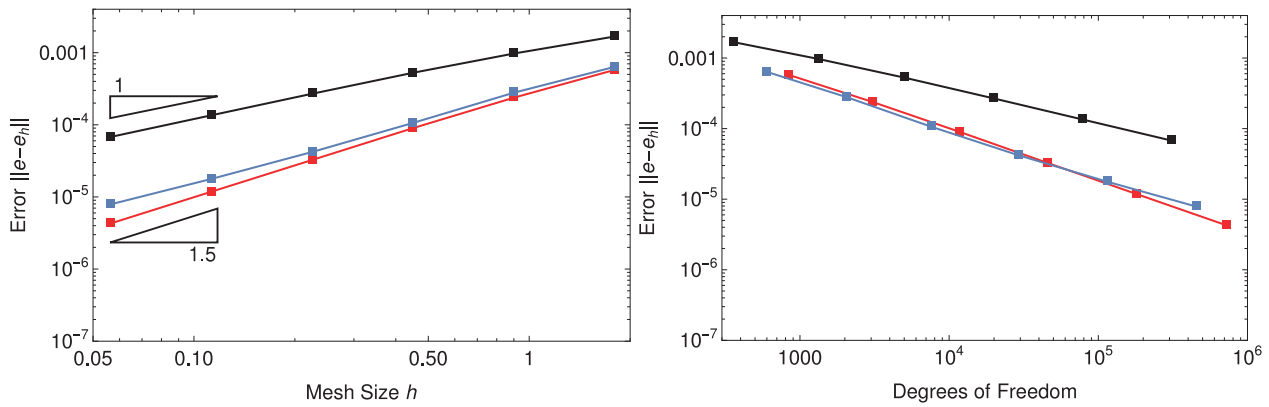


Fig. 16. Energy norm errors $\|e - e_h\|$ for different subdivision schemes: $\blacksquare$ $\mathbf{u}/\mathbf{p}/\mathbf{e}$, $\blacksquare$ $\mathbf{u}/\mathbf{p}$ and $\blacksquare$ AFR (reasonable, Fig. 13(a)).

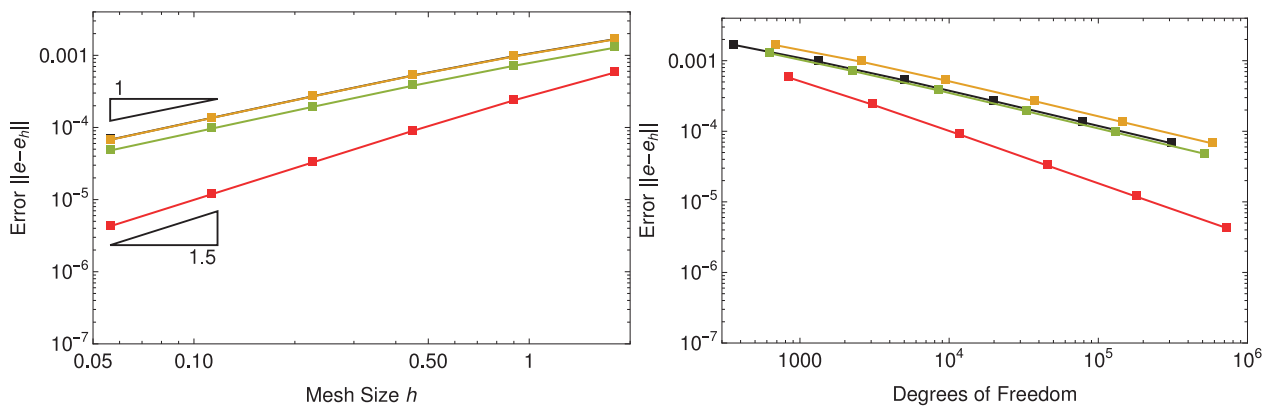


Fig. 17. Energy norm errors $\|e - e_h\|$ for different subdivision schemes: $\blacksquare$ $\mathbf{u}/\mathbf{p}/\mathbf{e}$, $\blacksquare$ $\mathbf{u}/\mathbf{p}$, $\blacksquare$ AFR (unreasonable, Fig. 13(b)) and $\blacksquare$ AFR (unreasonable, Fig. 13(c)).

the general convergence behavior. When displaying the error in the energy norm over the number of DOFs, it can be observed that the AFR technique performs as good as the $\mathbf{u}/\mathbf{p}/\mathbf{e}$ formulation, if a reasonable subdivision scheme is adopted.

Therefore, the initial findings of the local error analysis are confirmed: the results obtained using the AFR method are bounded by the $\mathbf{u}/\mathbf{p}$ and $\mathbf{u}/\mathbf{p}/\mathbf{e}$ formulations. Furthermore it is proven that the method converges towards the correct solution upon mesh refinement. The results also show that a reasonable choice of $\mathbf{u}/\mathbf{p}$ and $\mathbf{u}/\mathbf{p}/\mathbf{e}$ subdomains is important. The three-field formulation is most appropriate when large gradients are expected.

5.4. Friction stir welding application

The final example is selected to show that the AFR strategy can be employed for industrial applications. For the sake of simplicity, the geometric complexity of the actual pin-tool (e.g., threading and tilting, which is generally tackled with ALE or embedded formulations) is avoided. Only the large strain rate gradient in the thermo-mechanically affected zone is retained in the proposed analysis. Thus, a simpler cylindrical pin-tool is assumed to show the advantages and the accuracy of the AFR method.

Next, the following aspects of the analysis are described: (1) the description of the geometry and the corresponding mesh, (2) the loading and boundary conditions, (3) an overview of the temperature-dependent material parameters and (4) the obtained Finite Element results as well as an assessment of the influence of the AFR technique on the solution.

5.4.1. Geometry and mesh

The work pieces (the metal sheets to be welded) are modeled as flat hexahedral domains in an Eulerian framework, with a cylindrical hole, as shown in Fig. 18. The computational domain for the sheet is of size 100 mm $\times$ 100 mm with a thickness of 7 mm. The pin has a diameter of 6 mm and is inserted into the sheet to a depth of 6 mm. The tool shoulder has a diameter of 16 mm and is in contact with the sheet. The structured mesh consists of 122 425 nodes and 114 516 linear, hexahedral elements. When the $\mathbf{u}/\mathbf{p}$ formulation is used in the entire domain, the problem consists of approximately 500 000 DOFs. Using the $\mathbf{u}/\mathbf{p}/\mathbf{e}$ formulation everywhere, the total number of DOFs is approximately 1.2 million.

H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

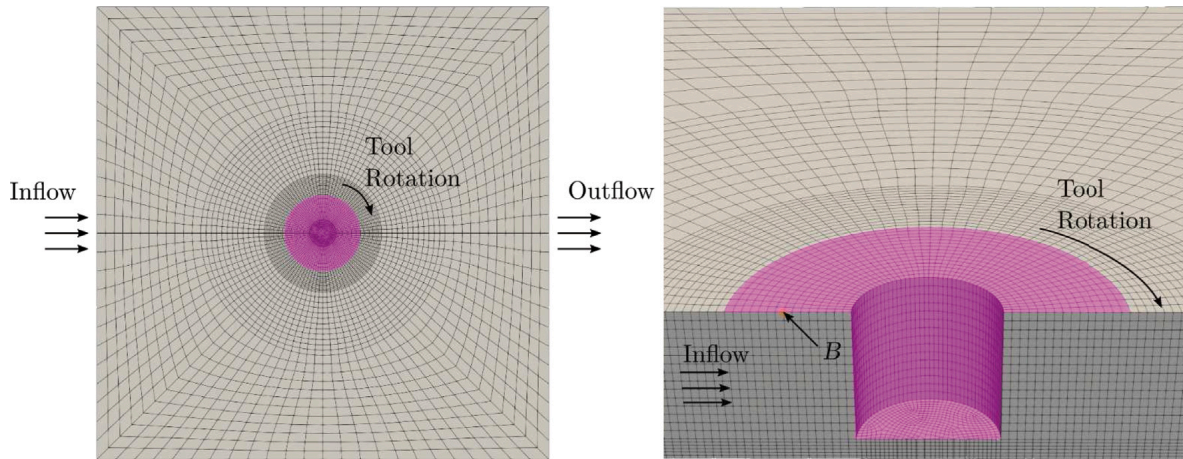


Fig. 18. Mesh of the three-dimensional FSW case which highlighted tool contact area.

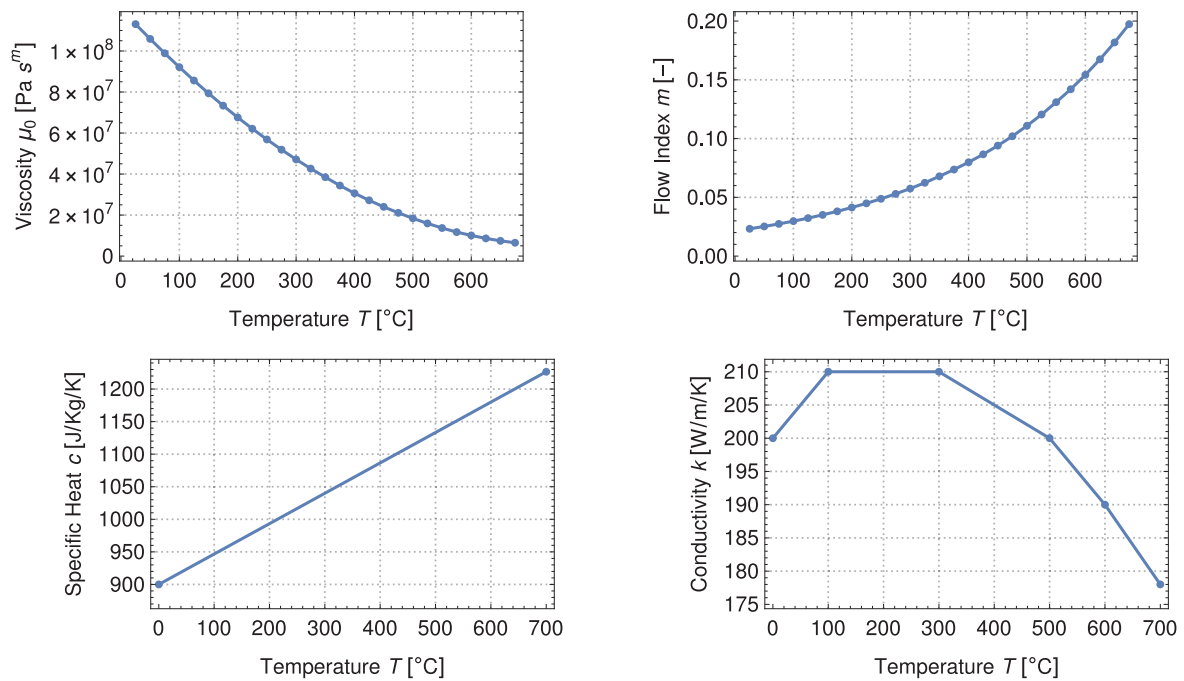


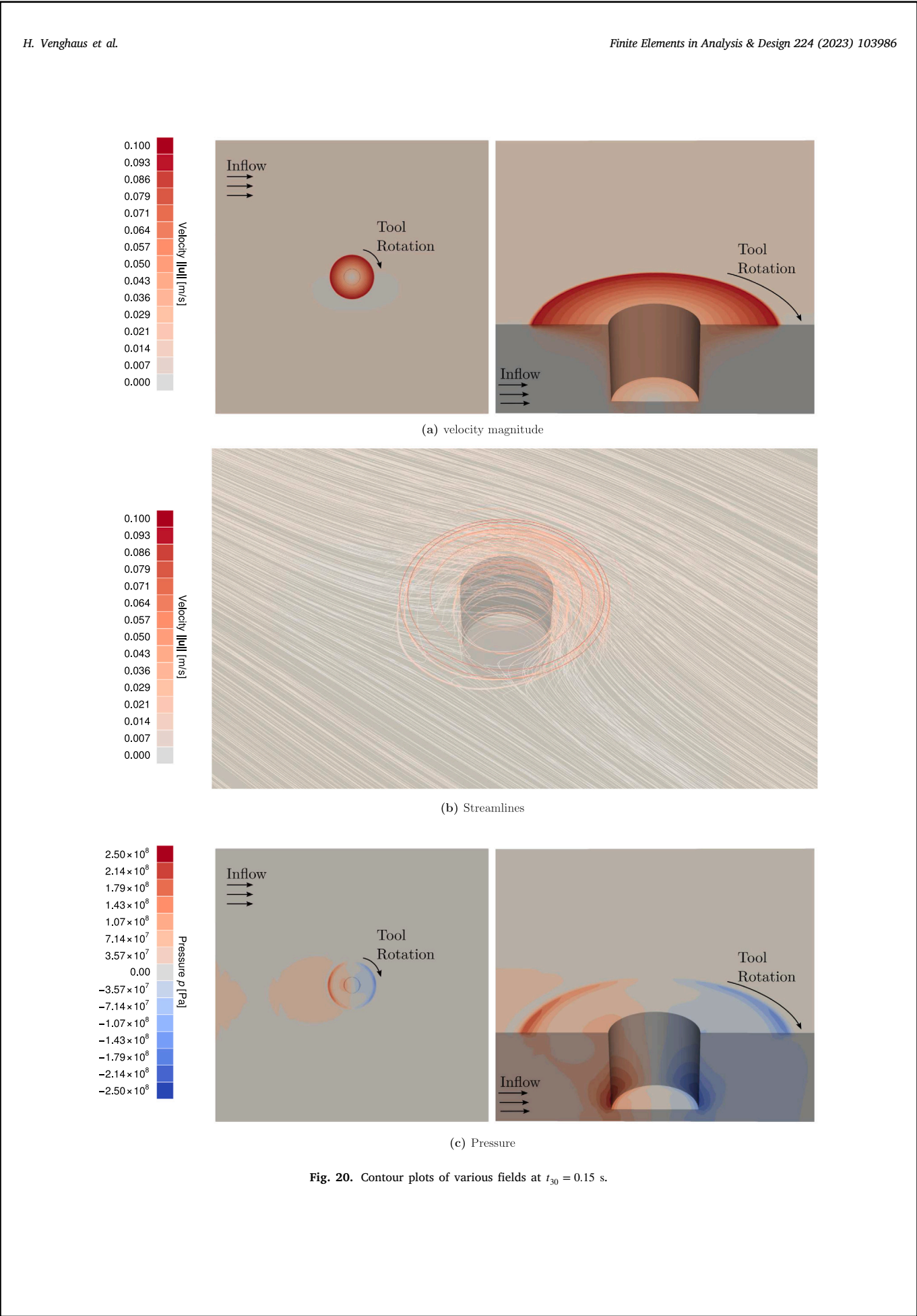
Fig. 19. Temperature dependency of mechanical and thermal material parameters.

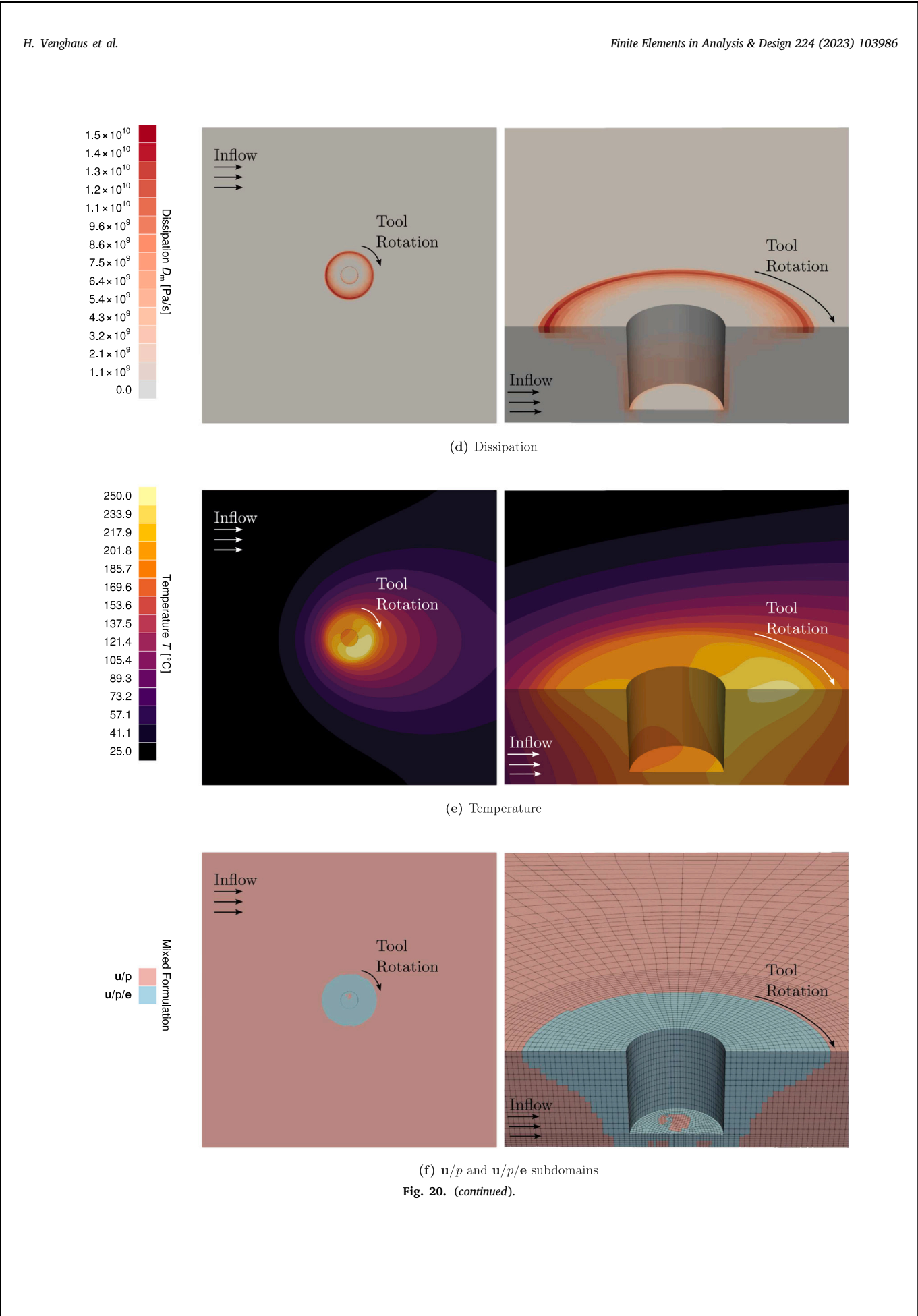
5.4.2. Loading and boundary conditions

The FSW process is simulated assuming a pin-tool rotation with an angular velocity of $f_{\text{tool}} = 100$ RPM. This rotation induces a maximum linear velocity on the outermost points of the shoulder of $u_{\text{max}} = 0.094248$ m/s. The pin-tool longitudinal movement (welding speed) is prescribed as a Dirichlet boundary condition for the sheet (relative movement). Therefore, at the front surface (labeled as *Inflow*) a constant inflow speed is prescribed at $u_{\text{in}} = 0.008333$ m/s. At the back surface (labeled *Outflow*) no Dirichlet boundary conditions are applied, allowing for the material to exit the domain. The material flow is assumed to be isochoric and slip wall boundary conditions are assumed for the rest of the surfaces defining the computational domain.

The interaction between the FSW pin-tool (not modeled) and the sheet is accounted for by prescribing its angular velocity at the (highlighted) contact surface. For the sake of simplicity, the friction law has been replaced by a fully-stick condition (Dirichlet condition). A study of different friction laws, to better capture the interaction between the pin-tool and the sheet can be found in [8].

The initial temperature as well as the environment temperature is set to 25 ° C. The heat conduction flow is set through the bottom surface into the back plate. The heat transfer coefficient by conduction is set to $h_{\text{cond}} = 2500$ W/m²/K.





H. Venghaus et al.

Finite Elements in Analysis & Design 224 (2023) 103986

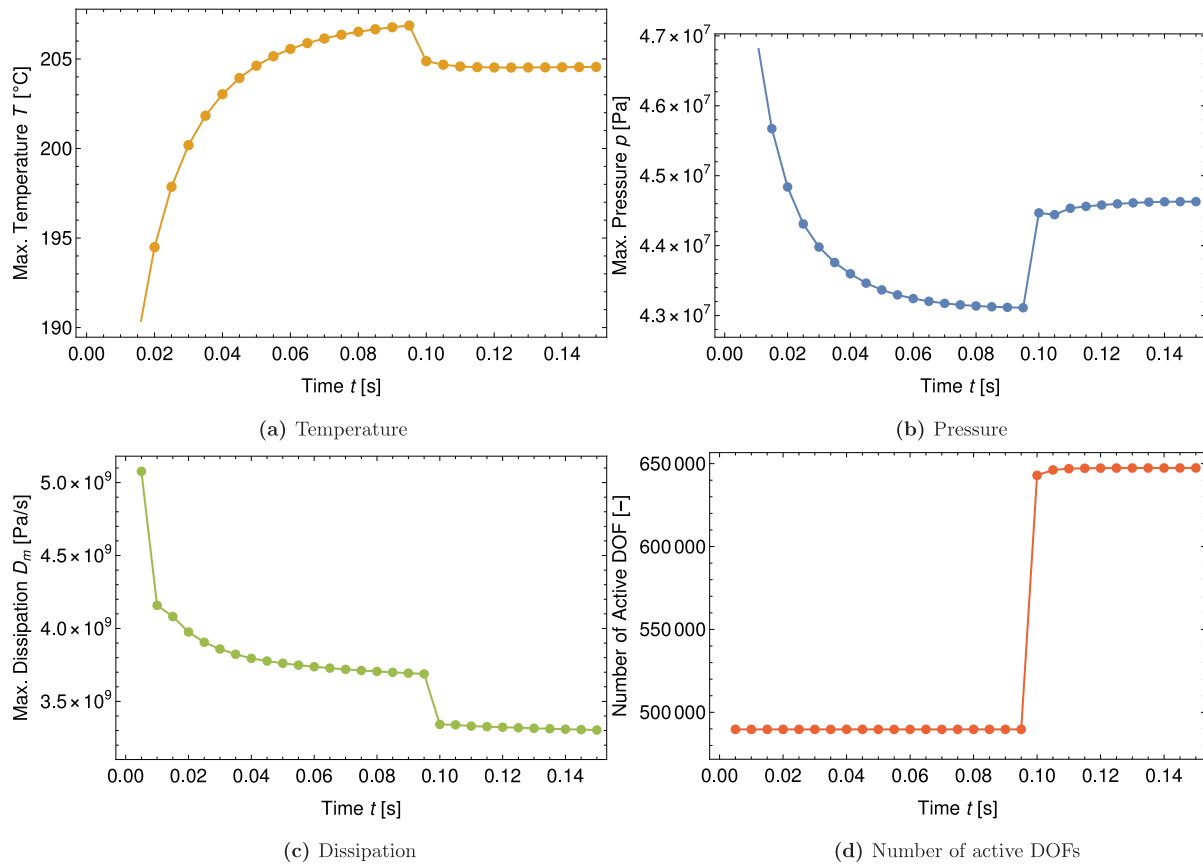


Fig. 21. Change of variables over time at point *B*.

All boundary conditions are applied at the beginning of the simulation (first time step) and are held constant throughout the entire simulation. In total, 30 time steps ($t_0 = 0$ s to $t_{30} = 0.15$ s) are computed. During the first 20 steps, the refinement criteria is disabled, allowing only for $\mathbf{u}/p$ elements to be used. Therefore, the steady state temperature field is reached as quickly as possible. At time step 20 ($t_{20} = 0.10$ s), the refinement criteria (based on a Zienkiewicz–Zhu (ZZ) error estimator with an error threshold of $1 \cdot 10^{-13}$) is enabled and all elements which exceed the specified error threshold become $\mathbf{u}/p/e$ elements. After the activation of the $\mathbf{u}/p/e$ elements, a new and more accurate steady state solution is reached.

5.4.3. Material law parameters

As introduced in Section 2.2, most material parameters are temperature-dependent as shown in Fig. 19. The fraction of the mechanical dissipation that is turned into heat (see Eq. (6)), is set to $\varphi = 0.8$. The density of the material is set to $\rho_0 = 2700$ kg/m³ and the artificial speed-up parameter (see Eq. (27)) to faster reach the steady state is set to $\alpha = 100$.

5.4.4. Results

The presentation and discussion of the results is split into two parts. First, the different solution fields at the last time step ($t_{30} = 0.15$ s) are shown as contour plots in Fig. 20. This allows to assess the results as a solution of the FSW analysis. Second, selected variables at a fixed point in the computational domain are plotted over time to show their sensitivity towards changing from $\mathbf{u}/p$ to $\mathbf{u}/p/e$ in parts of the domain.

The velocity field (Fig. 20a) is dominated by the Dirichlet boundary condition. The *shear-thinning* behavior of the applied material model can be clearly seen. Due to the high strain rates at the boundary layer, the viscosity decreases, causing a thinner boundary layer. Consequently, very high velocity gradients can be observed perpendicular to the top surface, further enhancing the shear-thinning effect. The same effect can be observed in Fig. 20b: only the streamlines closest to the pin and the shoulder are influenced by the tools rotation, whereas most other streamlines are relatively straight.

The pressure distribution (Fig. 20c) is almost symmetric and shows a spike on both the positive and the negative side where the shoulder of the tool ends. This is the expected result for the given boundary conditions. Using a proper friction model would yield more realistic results.

As expected, the contour plot of the dissipation (Fig. 20d) shows its highest values under the outermost part of the shoulder. Due to the shear-thinning behavior of the material model, the dissipation quickly drops by orders of magnitudes at a small distance to the tool-workpiece interface.

The temperature distribution (Fig. 20e) in the far-field shows a symmetric shape, while under the shoulder of the tool, the influence of the tools rotation can be seen. Furthermore, the influence of the convective term can be clearly seen.

The final plot (Fig. 20f) shows the u/p and $u/p/e$ subdomains. The area covered by $u/p/e$ elements is slightly larger than the area with significant dissipation, indicating that the refinement criterion and error threshold were well chosen.

Finally, Fig. 21 shows the evolution of different variables over time. The evolution of the temperature, pressure and dissipation (Figs. 21(a) to 21(c)) refers to point *B* (see Fig. 18). The fourth plot (Fig. 21(d)) is the number of active DOFs of the problem. In the first 20 steps ($t_0 = 0$ s to $t_{20} = 0.1$ s) only u/p elements are used. After activating the refinement criteria, the number of active DOFs increases by about 30%. The activation of the $u/p/e$ elements quickly leads to a new steady state.

It can be clearly seen from the plots, how temperature, pressure and dissipation change significantly after the activation of the $u/p/e$ elements—despite the already very fine mesh. As expected, the largest change can be observed for the dissipation, since the dissipation depends on both the strains and stresses. Thus, by employing the $u/p/e$ formulation, both fields are much more accurate.

Note that the new steady state is reached within 3 time steps. This emphasizes another advantage of the AFR method. It is possible to use the computationally cheaper u/p formulation to reach the steady state as quickly as possible and enhance the solution quality by locally activating the $u/p/e$ formulation.

6. Conclusions

In this paper, the proposed Adaptive Formulation Refinement method is used to define a computationally efficient framework for Friction Stir Welding simulations. The three-field $u/p/e$ formulation is used to enhance the strain and dissipation accuracy close the pin-tool, while the two-field u/p formulation is adopted to reduce the computational cost elsewhere. The use of linear Finite Elements is preserved for all unknowns by using a stabilized approach, the incompressible flow of the material is captured correctly and any remeshing is avoided. The AFR method is driven by the Zienkiewicz–Zhu a posteriori error estimator and the treatment of the interface between refined and unrefined subdomains is straightforward, due to the consistency of the variable fields and the lack of hanging nodes.

The global convergence study shows that the accuracy of the AFR method compares favorably to the $u/p/e$ formulation, while solving a system of equations with fewer active DOFs, thus resulting in faster build and solve times.

Finally, the developed method is applied to an FSW problem. First, the steady state of the problem is reached with a small computational cost by using the two-field formulation. Thereafter, the solution quality is enhanced in the highly dissipative area around the tool-sheet-interface, by switching to the three-field formulation according to the error estimator. Therefore the precision is very similar to the one of the three-field formulation but at a much lower computational cost.

CRediT authorship contribution statement

H. Venghaus: Methodology, Software, Validation, Formal Analysis, Writing – original draft, Visualization. **M. Chiumenti:** Conceptualization, Methodology, Writing – review & editing, Supervision. **J. Baiges:** Software, Validation, Writing – review & editing, Supervision. **D. Juhre:** Writing – review & editing, Supervision. **I. Castañar:** Methodology, Software, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

H. Venghaus acknowledges the support by Cimne, a Severo Ochoa Centre of Excellence (2019–2023) under the grant [CEX2018-000797-S], funded by the Research Agency of the Spanish State (AEI, [10.13039/501100011033]), the support from the PriMuS Project (Printing pattern based and MultiScale enhanced performance analysis of advanced Additive Manufacturing components, [PID2020-115575RB-I00]) through the Research Agency of the Spanish State (AEI, [10.13039/501100011033]) as well as the KYKLOS 4.0 project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement [No 872570]. I. Castanar gratefully acknowledges the support received from the Agència de Gestió d'Ajuts i de Recerca through the predoctoral FI grant 2019-FI-B-00649.

References

- [1] M. Chiumenti, et al., Numerical modeling of friction stir welding processes, *Comput. Methods Appl. Mech. Engrg.* 254 (2013) 353–369, <http://dx.doi.org/10.1016/j.cma.2012.09.013>.
- [2] N. Dialami, et al., Challenges in thermo-mechanical analysis of friction stir welding processes, *Arch. Comput. Methods Eng.* 24 (1) (2016) 189–225, <http://dx.doi.org/10.1007/s11831-015-9163-y>.
- [3] R. Padmanaban, V. Ratna Kishore, V. Balusamy, Numerical simulation of temperature distribution and material flow during friction stir welding of dissimilar aluminum alloys, *Procedia Eng.* 97 (2014) 854–863, <http://dx.doi.org/10.1016/j.proeng.2014.12.360>.
- [4] Saad B. Aziz, et al., Impact of friction stir welding (FSW) process parameters on thermal modeling and heat generation of aluminum alloy joints, *Acta Metall. Sin. (English Letters)* 29 (9) (2016) 869–883, <http://dx.doi.org/10.1007/s40195-016-0466-2>.
- [5] Rahul Kumar, V. Pancholi, R.P. Bharti, Material flow visualization and determination of strain rate during friction stir welding, *J. Mater. Process. Technol.* 255 (2018) 470–476, <http://dx.doi.org/10.1016/j.jmatprotec.2017.12.034>.
- [6] Narges Dialami, Miguel Cervera, Michele Chiumenti, Defect formation and material flow in friction stir welding, *Eur. J. Mech. A Solids* 80 (2020) 103912, <http://dx.doi.org/10.1016/j.euromechsol.2019.103912>.
- [7] C. Leitão, et al., Material flow in friction stir welding, *Microsc. Microanal.* 14 (S3) (2008) 87–90, <http://dx.doi.org/10.1017/s1431927608089472>.
- [8] N. Dialami, et al., Enhanced friction model for friction stir welding (FSW) analysis: Simulation and experimental validation, *Int. J. Mech. Sci.* 133 (2017) 555–567, <http://dx.doi.org/10.1016/j.ijmecsci.2017.09.022>.
- [9] D.G. Andrade, et al., Analysis of contact conditions and its influence on strain rate and temperature in friction stir welding, *Int. J. Mech. Sci.* 191 (2021) 106095, <http://dx.doi.org/10.1016/j.ijmecsci.2020.106095>.
- [10] Philippe Bussetta, et al., Comparison of a fluid and a solid approach for the numerical simulation of friction stir welding with a non-cylindrical pin, *Steel Res. Int.* 85 (6) (2013) 968–979, <http://dx.doi.org/10.1002/srin.201300182>.
- [11] Narges Dialami, Miguel Cervera, Michele Chiumenti, Numerical modelling of microstructure evolution in friction stir welding (FSW), *Metals* 8 (3) (2018) <http://dx.doi.org/10.3390/met8030183>, URL: <https://www.mdpi.com/2075-4701/8/3/183>.
- [12] M. Chiumenti, et al., Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation, *Finite Elem. Anal. Des.* 192 (2021) 103534, <http://dx.doi.org/10.1016/j.finel.2021.103534>.
- [13] T.J.R. Hughes, Equivalence of finite elements for nearly incompressible elasticity, *J. Appl. Mech.* 44 (1) (1977) 181–183, <http://dx.doi.org/10.1115/1.3423994>.
- [14] David S. Malkus, Thomas J.R. Hughes, Mixed finite element methods — Reduced and selective integration techniques: A unification of concepts, *Comput. Methods Appl. Mech. Engrg.* 15 (1) (1978) 63–81, [http://dx.doi.org/10.1016/0045-7825\(78\)90005-1](http://dx.doi.org/10.1016/0045-7825(78)90005-1).
- [15] Thomas J.R. Hughes, Generalization of selective integration procedures to anisotropic and nonlinear media, *Internat. J. Numer. Methods Engrg.* 15 (9) (1980) 1413–1418, <http://dx.doi.org/10.1002/nme.1620150914>.
- [16] E.A. Souza de Neto, et al., Design of simple low order finite elements for large strain analysis of nearly incompressible solids, *Int. J. Solids Struct.* 33 (20–22) (1996) 3277–3296, [http://dx.doi.org/10.1016/0020-7683\(95\)00259-6](http://dx.doi.org/10.1016/0020-7683(95)00259-6).
- [17] L. De Vivo, F. Marotti de Sciarra, The B-bar method and the limitation principles, *Int. J. Solids Struct.* 36 (34) (1999) 5177–5206, [http://dx.doi.org/10.1016/s0020-7683\(98\)00233-9](http://dx.doi.org/10.1016/s0020-7683(98)00233-9).
- [18] E.A. de Souza Neto, F.M. Andrade Pires, D.R.J. Owen, F-bar-based linear triangles and tetrahedra for finite strain analysis of nearly incompressible solids. Part I: formulation and benchmarking, *Internat. J. Numer. Methods Engrg.* 62 (3) (2005) 353–383, <http://dx.doi.org/10.1002/nme.1187>.
- [19] Y. Onishi, R. Iida, K. Amaya, F-bar aided edge-based smoothed finite element method using tetrahedral elements for finite deformation analysis of nearly incompressible solids, *Internat. J. Numer. Methods Engrg.* 109 (11) (2016) 1582–1606, <http://dx.doi.org/10.1002/nme.5337>.
- [20] J.C. Simo, M.S. Rifai, A class of mixed assumed strain methods and the method of incompatible modes, *Internat. J. Numer. Methods Engrg.* 29 (8) (1990) 1595–1638, <http://dx.doi.org/10.1002/nme.1620290802>.
- [21] E.A. Souza de Neto, et al., Remarks on the stability of enhanced strain elements in finite elasticity and elastoplasticity, *Commun. Numer. Methods. Eng.* 11 (11) (1995) 951–961, <http://dx.doi.org/10.1002/cnm.1640111109>.
- [22] Andreas Krischok, Christian Linder, On the enhancement of low-order mixed finite element methods for the large deformation analysis of diffusion in solids, *Internat. J. Numer. Methods Engrg.* 106 (4) (2015) 278–297, <http://dx.doi.org/10.1002/nme.5120>.
- [23] J.C. Simo, R.L. Taylor, K.S. Pister, Variational and projection methods for the volume constraint in finite deformation elasto-plasticity, *Comput. Methods Appl. Mech. Engrg.* 51 (1–3) (1985) 177–208, [http://dx.doi.org/10.1016/0045-7825\(85\)90033-7](http://dx.doi.org/10.1016/0045-7825(85)90033-7).
- [24] Peter Hansbo, Mats G. Larson, Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method, *Comput. Methods Appl. Mech. Engrg.* 191 (17–18) (2002) 1895–1908, [http://dx.doi.org/10.1016/s0045-7825\(01\)00358-9](http://dx.doi.org/10.1016/s0045-7825(01)00358-9).
- [25] Bernardo Cockburn, Dominik Schötzau, Jing Wang, Discontinuous Galerkin methods for incompressible elastic materials, *Comput. Methods Appl. Mech. Engrg.* 195 (25–28) (2006) 3184–3204, <http://dx.doi.org/10.1016/j.cma.2005.07.003>.
- [26] Adrian Lew, Alex Ten Eyck, Ramsharan Rangarajan, Some applications of discontinuous Galerkin methods in solid mechanics, in: *IUTAM Symposium on Theoretical, Computational and Modelling Aspects of Inelastic Media*, Springer Netherlands, 2008, pp. 227–236, http://dx.doi.org/10.1007/978-1-4020-9090-5_21.
- [27] Thomas J.R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, *Comput. Methods Appl. Mech. Engrg.* 127 (1–4) (1995) 387–401, [http://dx.doi.org/10.1016/0045-7825\(95\)00844-9](http://dx.doi.org/10.1016/0045-7825(95)00844-9).
- [28] Thomas J.R. Hughes, et al., The variational multiscale method—a paradigm for computational mechanics, *Comput. Methods Appl. Mech. Engrg.* 166 (1–2) (1998) 3–24, [http://dx.doi.org/10.1016/s0045-7825\(98\)00079-6](http://dx.doi.org/10.1016/s0045-7825(98)00079-6).
- [29] M. Chiumenti, et al., A stabilized formulation for incompressible elasticity using linear displacement and pressure interpolations, *Comput. Methods Appl. Mech. Engrg.* 191 (46) (2002) 5253–5264, [http://dx.doi.org/10.1016/s0045-7825\(02\)00443-7](http://dx.doi.org/10.1016/s0045-7825(02)00443-7).
- [30] Antoinette M. Maniatty, et al., Stabilized finite element method for viscoplastic flow: formulation and a simple progressive solution strategy, *Comput. Methods Appl. Mech. Engrg.* 190 (35–36) (2001) 4609–4625, [http://dx.doi.org/10.1016/s0045-7825\(00\)00346-7](http://dx.doi.org/10.1016/s0045-7825(00)00346-7).
- [31] Antoinette M. Maniatty, Yong Liu, Stabilized finite element method for viscoplastic flow: formulation with state variable evolution, *Internat. J. Numer. Methods Engrg.* 56 (2) (2002) 185–209, <http://dx.doi.org/10.1002/nme.554>.
- [32] M. Cervera, et al., Mixed linear/linear simplicial elements for incompressible elasticity and plasticity, *Comput. Methods Appl. Mech. Engrg.* 192 (49–50) (2003) 5249–5263, <http://dx.doi.org/10.1016/j.cma.2003.07.007>.
- [33] M. Cervera, et al., Softening, localization and stabilization: capture of discontinuous solutions in J2 plasticity, *Int. J. Numer. Anal. Methods Geomech.* 28 (5) (2004) 373–393, <http://dx.doi.org/10.1002/nag.341>.
- [34] M. Cervera, M. Chiumenti, C. Agelet de Saracibar, Shear band localization via local J2 continuum damage mechanics, *Comput. Methods Appl. Mech. Engrg.* 193 (9–11) (2004) 849–880, <http://dx.doi.org/10.1016/j.cma.2003.11.009>.
- [35] C. Agelet de Saracibar, et al., On the orthogonal subgrid scale pressure stabilization of finite deformation J2 plasticity, *Comput. Methods Appl. Mech. Engrg.* 195 (9–12) (2006) 1224–1251, <http://dx.doi.org/10.1016/j.cma.2005.04.007>.
- [36] Antoinette M. Maniatty, et al., Higher order stabilized finite element method for hyperelastic finite deformation, *Comput. Methods Appl. Mech. Engrg.* 191 (13–14) (2002) 1491–1503, [http://dx.doi.org/10.1016/s0045-7825\(01\)00335-8](http://dx.doi.org/10.1016/s0045-7825(01)00335-8).

- [37] Arif Masud, Timothy J. Truster, A framework for residual-based stabilization of incompressible finite elasticity: Stabilized formulations and $\bar{F}$ methods for linear triangles and tetrahedra, *Comput. Methods Appl. Mech. Engrg.* 267 (2013) 359–399, <http://dx.doi.org/10.1016/j.cma.2013.08.010>.
- [38] Guglielmo Scovazzi, et al., A simple, stable, and accurate linear tetrahedral finite element for transient, nearly, and fully incompressible solid dynamics: a dynamic variational multiscale approach, *Internat. J. Numer. Methods Engrg.* 106 (10) (2015) 799–839, <http://dx.doi.org/10.1002/nme.5138>.
- [39] S. Rossi, N. Abboud, G. Scovazzi, Implicit finite incompressible elastodynamics with linear finite elements: A stabilized method in rate form, *Comput. Methods Appl. Mech. Engrg.* 311 (2016) 208–249, <http://dx.doi.org/10.1016/j.cma.2016.07.015>.
- [40] Xianyi Zeng, et al., A dynamic variational multiscale method for viscoelasticity using linear tetrahedral elements, *Internat. J. Numer. Methods Engrg.* 112 (13) (2017) 1951–2003, <http://dx.doi.org/10.1002/nme.5591>.
- [41] Inocencio Castañar, Joan Baiges, Ramon Codina, A stabilized mixed finite element approximation for incompressible finite strain solid dynamics using a total Lagrangian formulation, *Comput. Methods Appl. Mech. Engrg.* 368 (2020) 113164, <http://dx.doi.org/10.1016/j.cma.2020.113164>.
- [42] H. Nguyen, p-Adaptive and Automatic Hp-Adaptive Finite Element Methods for Elliptic Partial Differential Equations (Ph.D. thesis), UC San Diego, 2010, URL: <https://escholarship.org/uc/item/8cv6d65z>.
- [43] Robert D. Russell, Adaptive mesh refinement, in: *Encyclopedia of Applied and Computational Mathematics*, Springer Berlin Heidelberg, 2015, pp. 23–25, http://dx.doi.org/10.1007/978-3-540-70529-1_341.
- [44] M. Cervera, M. Chiumenti, R. Codina, Mixed stabilized finite element methods in nonlinear solid mechanics. part I: Formulation, *Comput. Methods Appl. Mech. Engrg.* 199 (37–40) (2010) 2559–2570, <http://dx.doi.org/10.1016/j.cma.2010.04.006>.
- [45] G. Scovazzi, T. Song, X. Zeng, A velocity/stress mixed stabilized nodal finite element for elastodynamics: Analysis and computations with strongly and weakly enforced boundary conditions, *Comput. Methods Appl. Mech. Engrg.* 325 (2017) 532–576, <http://dx.doi.org/10.1016/j.cma.2017.07.018>.
- [46] M. Chiumenti, M. Cervera, R. Codina, A mixed three-field FE formulation for stress accurate analysis including the incompressible limit, *Comput. Methods Appl. Mech. Engrg.* 283 (2015) 1095–1116, <http://dx.doi.org/10.1016/j.cma.2014.08.004>.
- [47] M. Cervera, M. Chiumenti, R. Codina, Mixed stabilized finite element methods in nonlinear solid mechanics. part II: Strain localization, *Comput. Methods Appl. Mech. Engrg.* 199 (37–40) (2010) 2571–2589, <http://dx.doi.org/10.1016/j.cma.2010.04.005>.
- [48] M. Cervera, et al., Mixed stabilized finite element methods in nonlinear solid mechanics. part III: Compressible and incompressible plasticity, *Comput. Methods Appl. Mech. Engrg.* 285 (2015) 752–775, <http://dx.doi.org/10.1016/j.cma.2014.11.040>.
- [49] Carlos A. Moreira, et al., Accurate thermal-induced structural failure analysis under incompressible conditions, *Eng. Struct.* 261 (2022) 114213, <http://dx.doi.org/10.1016/j.engstruct.2022.114213>.
- [50] Inocencio Castañar, et al., Topological derivative-based topology optimization of incompressible structures using mixed formulations, *Comput. Methods Appl. Mech. Engrg.* 390 (2022) 114438, <http://dx.doi.org/10.1016/j.cma.2021.114438>.
- [51] Mark Ainsworth, J. Tinsley Oden, A posteriori error estimation in finite element analysis, *Comput. Methods Appl. Mech. Engrg.* 142 (1–2) (1997) 1–88, [http://dx.doi.org/10.1016/S0045-7825\(96\)01107-3](http://dx.doi.org/10.1016/S0045-7825(96)01107-3).
- [52] O.C. Zienkiewicz, J.Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, *Internat. J. Numer. Methods Engrg.* 24 (2) (1987) 337–357, <http://dx.doi.org/10.1002/nme.1620240206>.
- [53] R. Verfürth, A review of a posteriori error estimation techniques for elasticity problems, *Comput. Methods Appl. Mech. Engrg.* 176 (1–4) (1999) 419–440, [http://dx.doi.org/10.1016/S0045-7825\(98\)00347-8](http://dx.doi.org/10.1016/S0045-7825(98)00347-8).
- [54] Rüdiger Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Oxford University Press, 2013, <http://dx.doi.org/10.1093/acprof:oso/9780199679423.001.0001>.
- [55] T. Sheppard, D.S. Wright, Determination of flow stress: Part 1 constitutive equation for aluminium alloys at elevated temperatures, *Met. Technol.* 6 (1) (1979) 215–223, <http://dx.doi.org/10.1179/030716979803276264>.
- [56] Daniel Dall'Onder dos Santos, et al., Numerical approximations for flow of viscoplastic fluids in a lid-driven cavity, *J. Non-Newton. Fluid Mech.* 166 (12–13) (2011) 667–679, <http://dx.doi.org/10.1016/j.jnnfm.2011.03.004>.
- [57] Pierre J. Carreau, Rheological equations from molecular network theories, *Trans. Soc. Rheol.* 16 (1) (1972) 99–127, <http://dx.doi.org/10.1122/1.549276>.
- [58] F.H. Norton, The Creep of Steel At High Temperatures, in: *Bulletin*, McGraw-Hill book Company, Incorporated, 1929, URL: <https://books.google.de/books?id=vswxAAAAMAAJ>.
- [59] Nicholas John Hoff, Approximate analysis of structures in the presence of moderately large creep deformations, *Quart. Appl. Math.* 12 (1954) 49–55.
- [60] Franco Brezzi, Michel Fortin (Eds.), *Mixed and hybrid finite element methods*, Springer New York, 1991, <http://dx.doi.org/10.1007/978-1-4612-3172-1>.
- [61] Ramon Codina, Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods, *Comput. Methods Appl. Mech. Engrg.* 190 (13–14) (2000) 1579–1599, [http://dx.doi.org/10.1016/S0045-7825\(00\)00254-1](http://dx.doi.org/10.1016/S0045-7825(00)00254-1).
- [62] Ramon Codina, Jordi Blasco, Stabilized finite element method for the transient Navier–Stokes equations based on a pressure gradient projection, *Comput. Methods Appl. Mech. Engrg.* 182 (3–4) (2000) 277–300, [http://dx.doi.org/10.1016/S0045-7825\(99\)00194-2](http://dx.doi.org/10.1016/S0045-7825(99)00194-2).
- [63] Ramon Codina, Stabilized finite element approximation of transient incompressible flows using orthogonal subscales, *Comput. Methods Appl. Mech. Engrg.* 191 (39–40) (2002) 4295–4321, [http://dx.doi.org/10.1016/S0045-7825\(02\)00337-7](http://dx.doi.org/10.1016/S0045-7825(02)00337-7).
- [64] N. Dialami, et al., A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding, *Int. J. Mech. Sci.* 122 (2017) 215–227, <http://dx.doi.org/10.1016/j.ijmecsci.2016.12.016>.
- [65] C. Agelet de Saracibar, et al., Computational modeling and sub-grid scale stabilization of incompressibility and convection in the numerical simulation of friction stir welding processes, *Arch. Comput. Methods Eng.* 21 (1) (2014) 3–37, <http://dx.doi.org/10.1007/s11831-014-9094-z>.
- [66] Satish Balay, et al., *PETSc/TAO Users Manual*, Tech Report ANL-21/39 - Revision 3.16, Argonne National Laboratory, 2021.
- [67] J.E. Akin, An introduction to the fem and adaptive error analysis for engineering students, in: *Finite Element Analysis with Error Estimators*, Elsevier, 2005, <http://dx.doi.org/10.1016/B978-0-7506-6722-7.X5030-9>.

Publication 2

B



Contents lists available at ScienceDirect

Theoretical and Applied Fracture Mechanics

journal homepage: www.elsevier.com/locate/tafmec



Objectivity in quasi-brittle structural failure via adaptive formulation and mesh refinement

G.B. Barbat^{*}, M. Cervera, H. Venghaus, M. Chiumenti

International Center for Numerical Methods in Engineering (CIMNE), Technical University of Catalonia – BarcelonaTECH, Edificio C1, Campus Norte, Jordi Girona 1-3, 08034 Barcelona, Spain

ARTICLE INFO

Keywords:

Cracking
Structural failure
Mixed Finite Elements
Mesh adaptivity
Adaptive refinement

ABSTRACT

In this work, the combined implementation of *Adaptive Formulation Refinement* (AFR) and *Adaptive Mesh Refinement* (AMR) strategies is proposed for the efficient and mesh objective evaluation of quasi-brittle fracture.

A critical issue when computing localized structural failure is the spurious mesh bias dependence of the crack trajectories and collapse mechanisms that ensues when adopting the standard displacement-based finite element (FE) formulation. This problem is resolved by the introduction of the mixed strain/displacement FE formulation, which allows for local convergence, producing mesh objective results.

On the one hand, the proposed AFR approach allows to adaptively switch between the standard and mixed FE formulations during the numerical simulation depending on the level of accuracy required in different regions of the domain. Within this framework, the more accurate mixed formulation is activated only in the areas where the crack onset and evolution take place, while the standard formulation is adopted for the rest of the structure.

On the other hand, the adopted octree-based AMR strategy allows to adaptively refine only certain areas of the domain, also in function of the level of accuracy required. Using this methodology, it is possible to start the simulation with an initially relatively coarse mesh and adaptively introduce several levels of mesh refinement only in the regions of interest, where cracks appear and propagate. This approach guarantees the necessary mesh resolution to capture the fracture phenomenon accurately while further increasing the computational efficiency of the simulation. The subsequent adaptive coarsening of the FE mesh is also allowed.

Structural failure and fracture propagation are reproduced via a local isotropic damage model. For this reason, a refinement/coarsening criterion defined in function of the equivalent effective stress is conveniently adopted with both AFR and AMR strategies.

To assess the capabilities of the proposed method, a set of numerical simulations featuring benchmark problems and experiments is included. Results show the mesh objectivity and cost-efficiency of the approach combining AFR and AMR in fracture problems. The proposed model is able to reproduce experimental crack paths, collapse mechanisms and force–displacement curves with accuracy and without spurious mesh bias dependence.

The paper includes a comparative assessment of the performance of the proposed method and the standard FE formulation exclusively, showing the superior performance of the AFR methodology and its mesh objectivity. Convergence of the computed results when introducing different levels of AMR is also verified. Computational efficiency of the AFR and AMR techniques is specifically assessed. The combination of AFR and AMR approaches allows to provide accurate and mesh objective results in quasi-brittle fracture problems with very significant savings in computational cost.

1. Introduction

The numerical modelling of quasi-brittle fracture using the finite

element method has been intensively studied over the last five decades. The issue of the spurious dependency of the computed crack paths with the FE mesh employed when adopting the standard displacement-based

^{*} Corresponding author.

E-mail addresses: gabriel.barbat@upc.edu (G.B. Barbat), miguel.cervera@upc.edu (M. Cervera), henning.venghaus@upc.edu (H. Venghaus), michele.chiumenti@upc.edu (M. Chiumenti).

<https://doi.org/10.1016/j.tafmec.2022.103646>

Received 20 July 2022; Received in revised form 6 October 2022; Accepted 18 October 2022

Available online 28 October 2022

0167-8442/© 2022 Elsevier Ltd. All rights reserved.

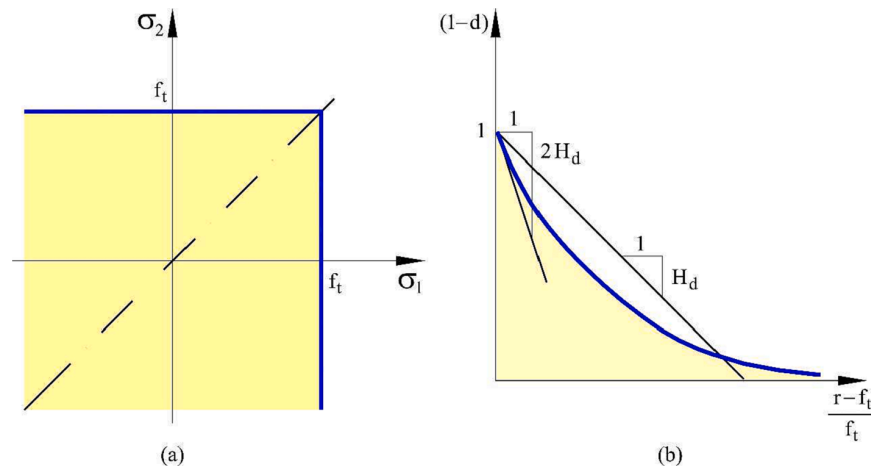


Fig. 1. Damage model: (a) Rankine damage surface and (b) softening function.

formulation has been reported since the 1980s [1,2,3,4]. This has resulted in the adoption of many different alternative techniques to model fracture, both at the continuous and discontinuous levels, trying to overcome this issue. A comparative review of alternative methodologies can be found in references [5,6].

The spurious mesh dependence that takes place when using the standard FE formulation from solid mechanics is caused by the lack of local convergence of stresses and strains. Depending on the cracking criterion, the trajectory of the fractures is governed by the stress fields present at the tip of the crack. These are inaccurately computed with standard FEs, resulting in unreliable and spuriously mesh biased results. This idea is underlined in references [6,7,8,9].

In view of this, in previous works [7,8,10,11,12,13,14,15], the authors have proposed the use of the strain/displacement ϵ/u finite element formulation to compute fracture with enhanced accuracy and mesh objectivity. In this FE formulation, the strain field is treated as an additional primary unknown of the problem and is interpolated independently from the displacements, rather than being computed by local differentiation at element level. This allows to improve the order of convergence and the accuracy of the computed strains and stresses. Specifically, it allows to achieve local convergence of the discrete strain and stress fields in the region of the crack tip, this being a critical feature to produce mesh bias objective results.

For the same reason, by introducing the volumetric/deviatoric split into the solid mechanics problem, the 3-field displacement/deviatoric-strain/pressure $u/e/p$ FE formulation has been proposed, showing to produce more accurate and reliable computations with respect to the classical u/p FE in nonlinear problems under incompressible conditions [16,17,18,19,20,21].

Therefore, on the one hand, despite the issue of producing spuriously mesh dependent results, the standard FE formulation is used in many applications due to its simplicity and affordable computational cost, which makes it convenient for large-scale analyses. On the other hand, the mixed ϵ/u FE formulation has shown to be able of providing numerical results with mesh objectivity and enhanced accuracy, at the expense of treating the strains as additional nodal unknowns of the problem and introducing a relatively larger computational cost in the FE computation.

Notwithstanding, in the specific problem of quasi-brittle fracture, structural failure occurs in the form of a few cracks appearing and evolving in an area representing a reduced part of the whole domain. Both the standard and mixed formulations are compatible and the use of the more accurate mixed formulation is required only in the small part of the total domain where crack propagation occurs while in the rest of the structure the less costly standard formulation may be adopted.

In view of this, in the present work the *Adaptive Formulation Refinement* (AFR) strategy is implemented to adaptively switch between the standard and mixed ϵ/u FE formulations in different regions of the domain depending on the required accuracy, so that an efficient computation framework is defined. The use of isotropic damage models to represent the degradation of the material allows to propose a criterion established in terms of the equivalent effective stress to decide on toggling between the standard and mixed formulations. The AFR methodology was originally proposed in reference [22] for selectively activating the u/p and $u/e/p$ mixed formulations in the simulation of friction stir welding (FSW) processes.

In the present work, the use of the AFR strategy is proposed in conjunction with an *Adaptive Mesh Refinement* (AMR) framework. AMR has been often employed in alternative models for fracture [23,24,25,26,27]. The adopted approach uses an octree-based method [28,29,30,31] to further increase the efficiency of the numerical model by adaptively refining only the areas of the domain where crack propagation occurs. This allows to reproduce the phenomenon of fracture with proper FE resolution while maintaining a coarser mesh elsewhere. The method includes the adaptive refinement and coarsening of the FE mesh according to a criterion also defined in terms of the equivalent effective stress.

The objectives of the present work are: (1) to establish the necessary features required in the implementation of the approach for quasi-brittle fracture based on the AFR and AMR strategies, (2) to demonstrate the capabilities of the proposed model in structural failure problems, (3) to show the ability of the model in reproducing experimental results in terms of collapse mechanisms, force–displacement curves and crack trajectories with precision, (4) to verify the capacity of the proposed approach combining the AFR and AMR methodologies to produce mesh objective results, (5) to assess the performance of the method with regard to accuracy and computational efficiency.

The outline of the paper is the following. In Section 2 the mixed strain/displacement ϵ/u finite element formulation used to solve the structural problem is introduced, including the strong and weak continuous forms as well as the corresponding discrete FE approximation. The isotropic damage constitutive law considered to model cracking is presented as well. Section 3 details the AFR (Section 3.1) and AMR (Section 3.2) methodologies adopted. Section 4 includes the numerical simulations of structural failure performed to assess the capabilities of the proposed model. Section 5 closes the paper with the conclusions of the work.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

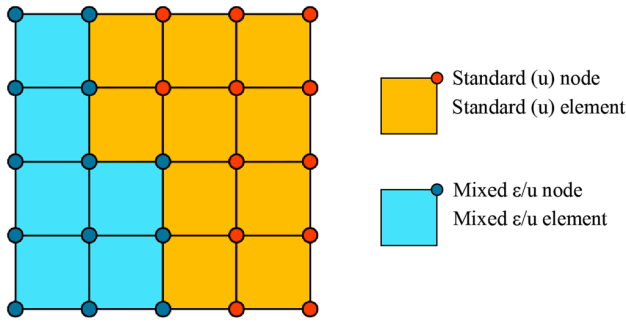


Fig. 2. Finite element mesh combining standard and mixed formulations.

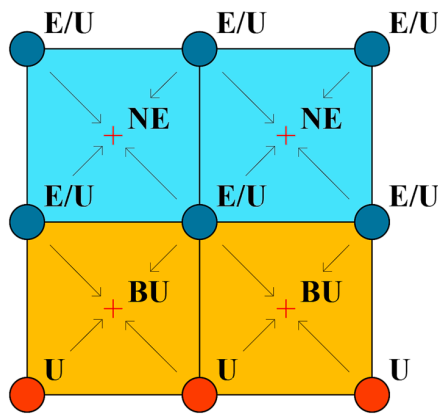


Fig. 3. Computation of strains at integration points (represented with red crosses). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Finite element formulation

Many different methods for the FE computation of quasi-brittle fracture have been proposed in the literature. By far, the most employed one, both by academics and practitioners, is the *crack band model*, which relies on the use of the standard displacement-based FE formulation for continuum solid mechanics (CSM) problems. This approach, proposed by Rashid in 1968 [32], is applied together with a constitutive law featuring softening such as damage or plasticity to represent material failure at local level. In 1983, Bazant and Oh [33] introduced the objective evaluation of the softening behavior of the material with respect to the FE size, allowing to obtain results independent of the FE mesh size when computing fracture with the crack band model.

Regrettably, the trajectories of the cracks computed with the standard displacement-based FE formulation are spuriously dependent on the orientation of the mesh employed for the analysis. This well-known issue since the 1980's [1,2,3,4] renders the method useless in practice.

Although not generally recognized, this is caused by the lack of local convergence of stresses and strains of the standard FE formulation [7,8]. Near the vicinity of the tip of the crack, the stress and strain fields which are locally used to determine the path of the fracture are computed inaccurately, resulting in the spurious mesh bias dependence of the results.

The use of the mixed strain/displacement ϵ/u FE formulation for solid mechanics has been proposed by Cervera and coworkers [7,8,10,11,12,13,14,15] as an effective remedy. In this approach, the strains are treated as additional primary unknowns of the problem, together with the displacements. The mixed formulation has proven to

be locally convergent in terms of strains and stresses [7,8] and, consequently, to produce mesh-bias independent results in cracking problems.

Previous studies analyzing the comparative performance of the standard and mixed ϵ/u formulations have extensively shown, on the one hand, the inability of the standard FEs of delivering results without spurious mesh dependence on the FE orientation and, on the other hand, the capacity of the mixed FEs for consistently producing reliable mesh objective computations [13,15]. Also, it has been demonstrated that convergence of the computed stresses, strains and mechanical dissipation is guaranteed with mixed FEs [18].

Notwithstanding, the mixed ϵ/u FE formulation introduces the strain field as an additional primary unknown of the solid mechanics problem being solved, together with the displacements, resulting in a higher cost of the numerical simulations. In the present work the efficiency of the mixed FEM is increased by adaptively activating the mixed formulation only in the regions where local accuracy is required in order to obtain convergent and mesh objective results, while using the standard formulation elsewhere. Computational efficiency of the mixed FEM is further increased in the present work with the adoption of the Adaptive Mesh Refinement approach.

The proposed approach preserves the local format of the problem. Fracture propagation is computed according to the local values of the stress field at the tip of the developing crack, without introducing gradient or higher order terms in the constitutive law or regularizing terms in the variational form. Note that, as it is pointed out in reference [6], crack models regularized at continuous level require the adoption of very fine meshes. This results in an important increase of the computational cost involved in the calculations when compared to the crack band model, much larger than the one ensuing from the use of the mixed ϵ/u formulation [6]. For this reason in regularized approaches, such as phase-field models, adaptive mesh refinement capabilities are often adopted [23,24,25,26,27]. The issue of computational cost has also been addressed in reference [34] by computing the nodal phase-field variable only in a predefined reduced sub-domain where cracking occurs rather than in the whole structure. An appraisal of the computational costs involved in the alternative methods proposed for the evaluation of localized structural failure can be found in the comparative study of reference [6].

2.1. Continuous formulation

In the rest of this section, the mixed strain/displacement ϵ/u FE formulation employed for the mesh objective evaluation of quasi-brittle failure is presented. For additional details the interested reader is referred to references [7,8,11,12,14,35,36,37].

The displacements u and the strains ϵ are adopted as primary unknowns of the nonlinear solid mechanics problem. The compatibility equation relates these two fields in the following way:

$$\epsilon = Su \quad (1)$$

S being the differential symmetric gradient operator.

Also, the Cauchy momentum equation associates the stresses σ and the body forces f

$$S^T \sigma + f = 0 \quad (2)$$

S^T being the differential divergence operator, adjoint to the S in Eq. (1).

In addition, the stresses and strains are connected through the constitutive law, established in secant form as

$$\sigma = D\epsilon \quad (3)$$

where $D = D(d)$ is the (nonlinear) secant constitutive matrix, expressed in function of the damage index, d . D needs to be symmetric and positive semidefinite from thermodynamic considerations.

From the three previous equations, the strong form of the problem is

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

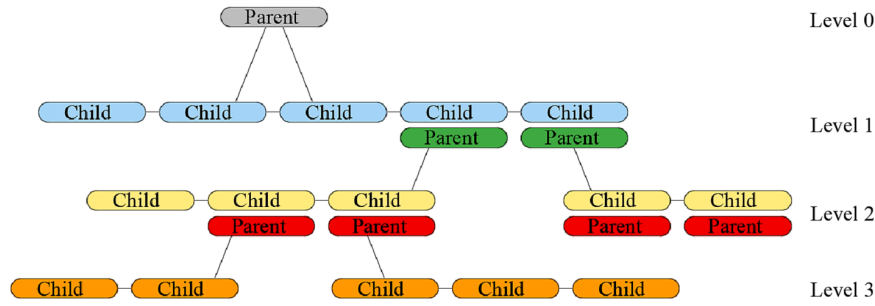


Fig. 4. Finite element hierarchy in the octree-based AMR strategy.

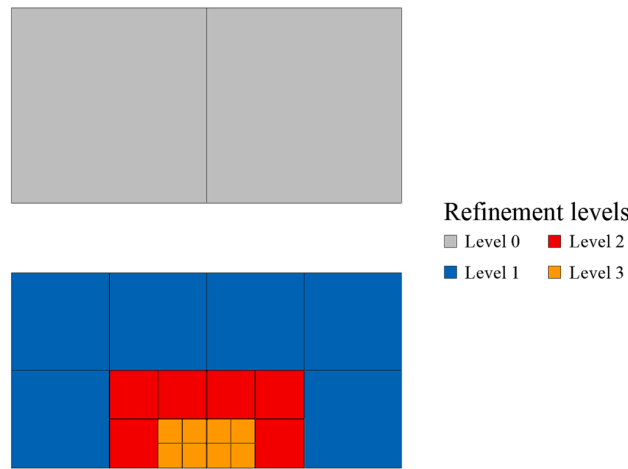


Fig. 5. Depiction of the octree-based refinement process in a finite element.

obtained as follows: firstly, Eq. (1) is pre-multiplied by $\mathbf{D}$, and secondly, Eq. (3) is introduced into Eq. (2)

$$-\mathbf{D}\boldsymbol{\varepsilon} + \mathbf{D}\mathbf{S}\mathbf{u} = 0 \quad (4)$$

$$\mathbf{S}^T(\mathbf{D}\boldsymbol{\varepsilon}) + \mathbf{f} = 0 \quad (5)$$

Then the corresponding weak form is derived by: (i) multiplying Eqs. (4) and (5) by vectors of virtual strains and virtual displacements respectively, (ii) integrating both equations over the spatial domain and (iii) applying the Divergence Theorem to the first term of Eq. (5). This results in the following system of equations:

$$-\int_{\Omega} \delta\boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon} d\Omega + \int_{\Omega} \delta\boldsymbol{\varepsilon}^T \mathbf{D}\mathbf{S}\mathbf{u} d\Omega = 0 \quad \forall \delta\boldsymbol{\varepsilon} \quad (6)$$

$$\int_{\Omega} (\mathbf{S}\delta\mathbf{u})^T (\mathbf{D}\boldsymbol{\varepsilon}) d\Omega = \int_{\Omega} \delta\mathbf{u}^T \mathbf{f} d\Omega + \int_{\Gamma_t} \delta\mathbf{u}^T \bar{\mathbf{t}} d\Gamma \quad \forall \delta\mathbf{u} \quad (7)$$

Note that the boundary Γ of the domain is divided in two parts, Γ_u and Γ_t , corresponding to the Dirichlet and Newmann boundary conditions.

The variational form of the problem consists in finding the strain and displacement fields $\boldsymbol{\varepsilon}$ and $\mathbf{u}$ complying with the system of Eqs. (6)-(7) and conforming to the boundary condition $\mathbf{u} = 0$ in Γ_u given the arbitrary virtual displacement vector $\delta\mathbf{u}$, which is also null on Γ_u , and the arbitrary virtual strain vector $\delta\boldsymbol{\varepsilon}$. Note as well that the problem is symmetric.

2.2. FE approximation

Once the spatial domain Ω is discretized into nonintersecting finite elements Ω_e , such that $\Omega = \cup \Omega_e$, the displacement $\mathbf{u}$ and the strain $\boldsymbol{\varepsilon}$ are approximated using independent interpolations $\hat{\mathbf{u}}$ and $\hat{\boldsymbol{\varepsilon}}$.

$$\mathbf{u} \cong \hat{\mathbf{u}} = \mathbf{N}_u \mathbf{U} \quad (8)$$

$$\boldsymbol{\varepsilon} \cong \hat{\boldsymbol{\varepsilon}} = \mathbf{N}_\varepsilon \mathbf{E} \quad (9)$$

where $\mathbf{U}$ and $\mathbf{E}$ are vectors containing the nodal values of the displacements and the strains. $\mathbf{N}_u$ and $\mathbf{N}_\varepsilon$ are the matrices containing the interpolation functions employed in the FE approximation.

When adopting a mixed FE, the interpolation functions employed need to fulfill the inf-sup stability condition to avoid spurious oscillations in the computed solution [38,39,40]. This condition is not satisfied when equal order linear interpolation functions are considered for $\mathbf{N}_u$ and $\mathbf{N}_\varepsilon$ [39,41,42]. In order to use linear interpolation functions for both fields while avoiding spurious oscillations in the computed solution, a stabilization procedure is introduced. The proposed method follows the Variational Multiscale Stabilization approach [43,44,45,46,47].

The procedure is to substitute the strain field approximation in Eq. (9) with the following one

$$\boldsymbol{\varepsilon} \cong \hat{\boldsymbol{\varepsilon}} = \mathbf{N}_\varepsilon \mathbf{E} + \tau_\varepsilon (\mathbf{B}_u \mathbf{U} - \mathbf{N}_\varepsilon \mathbf{E}) = (1 - \tau_\varepsilon) \mathbf{N}_\varepsilon \mathbf{E} + \tau_\varepsilon \mathbf{B}_u \mathbf{U} \quad (10)$$

where τ_ε is a stabilization parameter such that $0 \leq \tau_\varepsilon \leq 1$. $\mathbf{B}_u$ is the standard compatibility matrix, defined as $\mathbf{B}_u = \mathbf{S}\mathbf{N}_u$. Using $\tau_\varepsilon = 0$ corresponds to adopting the strain interpolation of the non-stabilized mixed problem, while the case $\tau_\varepsilon = 1$ recovers the strain interpolation of the standard displacement-based formulation from solid mechanics.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

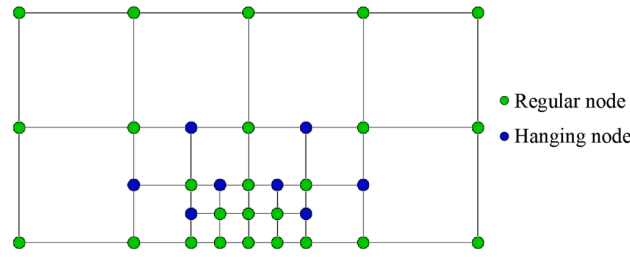


Fig. 6. Regular and hanging nodes in the FE mesh.

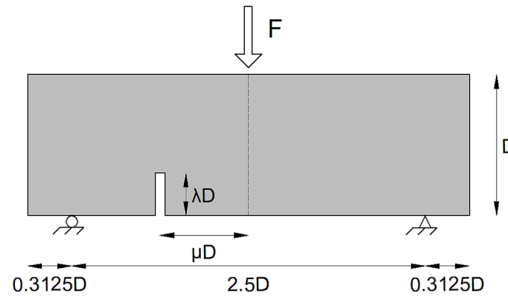


Fig. 7. Geometry of the Garcia-Alvarez beams.

References [7,8,47] have shown that the optimal convergence rate of the problem is attained when considering $\tau_\epsilon = c_\epsilon h/L$, where $c_\epsilon = \mathcal{O}(1)$ is an arbitrary constant, L is a reference size of the structure and h is the FE size. In the present work, $c_\epsilon = 1$ is used. When adaptive mesh refinement is introduced according to the procedure described in the following sections, the FE size h used to compute τ_ϵ is adjusted accordingly.

Also, taking into consideration the definition of the strain field in Eq. (10), the value of the crack bandwidth b becomes:

$$b = (1 - \tau_\epsilon)2h + \tau_\epsilon h = (2 - \tau_\epsilon)h \quad (11)$$

By introducing the interpolation function proposed in Eqs. (8) and (10) into the weak form of the problem (6)-(7), the following algebraic system of equations results

$$\begin{bmatrix} -M_\tau & G_\tau \\ G_\tau^T & K_\tau \end{bmatrix} \begin{bmatrix} E \\ U \end{bmatrix} = \begin{bmatrix} 0 \\ F \end{bmatrix} \quad (12)$$

where $[E \ U]^T$ is the vector containing the nodal values of the unknowns of the problem, strains and displacements, and $M_\tau = (1 - \tau_\epsilon)M$, $G_\tau = (1 - \tau_\epsilon)G$ and $K_\tau = \tau_\epsilon K$. M is a mass-like projection matrix, G is the discrete gradient matrix, K is a stiffness-like matrix and F is the vector of external nodal forces:

$$M = \int_{\Omega} N_\epsilon^T \mathbf{D} N_\epsilon d\Omega \quad (13)$$

$$G = \int_{\Omega} N_\epsilon^T \mathbf{D} \mathbf{B}_u d\Omega \quad (14)$$

$$K = \int_{\Omega} \mathbf{B}_u^T \mathbf{D} \mathbf{B}_u d\Omega \quad (15)$$

$$F = \int_{\Omega} N_u^T \mathbf{f} d\Omega + \int_{\Gamma_t} N_u^T \bar{t} d\Gamma \quad (16)$$

The use of the mixed strain/displacement FE formulation in CSM problems presents the following advantages:

- It does not require any specific FE interpolation and can be used with triangles, quadrilaterals, tetrahedra, hexahedra and prisms [11].

- It does not need any specific development for implementation in 3D [11].
- It can be employed with any suitable constitutive behavior [12,14,35].

When used in problems involving quasi-brittle fracture:

- Previous works have shown its applicability with isotropic and orthotropic damage models including crack closure-reopening effects and irreversible straining [12,14]. Plasticity constitutive laws have also been adopted within this FE formulation [35,48,49,50].
- It can accurately reproduce the structural size effect phenomenon [15].
- It can naturally reproduce branching and intersecting cracks without any additional specific developments [6].

2.3. Damage constitutive laws

In this work, cracks are modelled using an isotropic continuum damage model, a reformulation of the original crack band model. Alternatively, orthotropic damage models can be adopted [12,14].

Two versions of the constitutive law are proposed, one featuring the Rankine damage criterion and the other one with the Drucker-Prager damage criterion.

For an isotropic damage model, the constitutive law in Eq. (3) is expressed as

$$\sigma = \mathbf{D}(d)\epsilon = (1 - d)\mathbf{D}_0\epsilon \quad (17)$$

where $\mathbf{D} = \mathbf{D}(d)$ is the secant constitutive matrix, written in terms of the internal scalar damage variable d which represents the degradation state of the material. $\mathbf{D}_0$ is the initial elastic constitutive matrix of the undamaged material, which is defined for initially isotropic elastic materials as a function of the initial values of Young's modulus E and Poisson's ratio ν . $\mathbf{D}$ is symmetric and positive semidefinite for $0 \leq d \leq 1$.

The damage criterion $\mathbb{F}$ is introduced as

$$\mathbb{F}(\bar{\sigma}_{eq}, r) = \bar{\sigma}_{eq}(\bar{\sigma}) - r = 0 \quad (18)$$

where the effective stress $\bar{\sigma}$ is introduced as $\bar{\sigma} = \mathbf{D}_0\epsilon$, following the

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

Table 1
Material parameters of the Garcia-Alvarez beams.

Young's Modulus	33.8·10 ⁹ Pa
Poisson's Ratio	0.2
Tensile Strength	3.5·10 ⁶ Pa
Tensile Fracture Energy	80 J/m ²

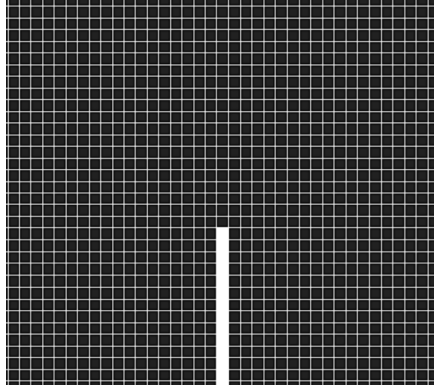


Fig. 8. Detail of the initial mesh used for the Garcia-Alvarez beams in the area of the notch.

hypothesis of strain equivalence, the equivalent effective stress $\bar{\sigma}_{eq}$ is defined in function of the damage criterion considered and r is the current damage threshold.

For the Rankine damage criterion, whose surface is plotted in Fig. 1a, the equivalent effective stress $\bar{\sigma}_{eq}$ is computed in function of $\bar{\sigma}$ as

$$\bar{\sigma}_{eq}(\bar{\sigma}) = \langle \bar{\sigma}_1 \rangle \quad (19)$$

where $\bar{\sigma}_1$ is the major principal effective stress and $\langle \cdot \rangle$ are the Macaulay brackets, such that $\langle x \rangle = x$ if $x \geq 0$ and $\langle x \rangle = 0$ if $x < 0$.

Alternatively, for the Drucker-Prager damage criterion

$$\bar{\sigma}_{eq}(\bar{\sigma}) = \frac{\sqrt{3J_2} + \alpha I_1}{1 + \alpha} \quad (20)$$

where I_1 and J_2 are defined in function of $\bar{\sigma}$ as the first and second effective stress invariants and α is calculated in function of the tensile and compressive strengths of the material f_t and f_c

$$\alpha = \left(\frac{f_c}{f_t} - 1 \right) / \left(\frac{f_c}{f_t} + 1 \right)$$

Other damage criteria may be readily used, as shown in references [12,14].

The initial value of the damage threshold r is the tensile strength of the material f_t . Its current value at time t is determined according to the Kuhn-Tucker optimality and consistency conditions, guaranteeing the positiveness of the dissipation and the irreversibility of damage.

$$r = \max [f_t, \max \bar{\sigma}_{eq}(\hat{t})] \quad \hat{t} \in [0, t] \quad (21)$$

In the present work it is considered that the internal damage variable d follows an exponential softening behavior, as shown in Fig. 1b:

$$d = 1 - \frac{f_t}{r} \exp \left(-2H_d \frac{(r - f_t)}{f_t} \right) \quad (22)$$

where H_d is a softening parameter controlling the rate of degradation of the material.

Alternative softening functions may be used if deemed necessary. They are required to monotonically decrease from 1 to 0 as the damage threshold r evolves from f_t to infinity. Also, the objective definition of the mechanical dissipation needs to be established for each function.

Energy conservation considerations link the softening parameter H_d

in Eq. (22) to material properties via Irwin's material length $\mathcal{L} = EG_f / f_t^2$ and the crack bandwidth b .

$$H_d = \frac{b}{2\mathcal{L} - b} \quad (23)$$

In this way, the proposed model is able to provide results independent of the crack bandwidth, guaranteeing the objective energy dissipation during the cracking process. The crack bandwidth is computed as in Eq. (11), which yields $b = h$ when using the standard FE formulation ($\tau_e = 1$).

3. Adaptivity in quasi-brittle structural failure computations

In this section, the approaches employed for introducing the adaptive mesh refinement (AMR) technique and the adaptive FE formulation refinement (AFR) into the FE computations are presented. They are both implemented with the aim of further increasing the computational efficiency of the proposed mixed e/u FE method for computing quasi-brittle fracture while preserving its capacity for producing convergent results with enhanced accuracy and mesh objectivity.

3.1. Adaptive formulation refinement

The present work uses the adaptive application of the standard and mixed strain/displacement FE formulations, as depicted in Fig. 2.

References [36,51] underline the fact that both formulations are compatible. In order to ensure the consistency of the AFR approach in the classical Rayleigh-Ritz sense, the continuity of the interpolated displacement field across element edges needs to be ensured. This requirement is fulfilled when the same interpolation functions are adopted for the displacements for both the standard and mixed formulations. For further details references [36,51] are recommended.

For this reason, the less costly standard displacement-based formulation may be adaptively switched with the enhanced accuracy mixed formulation only when and where necessary. In the following, the main features of the adopted AFR strategy are presented. For additional details, the original reference [22] introducing the method is recommended.

3.1.1. Strain and stress computations

Strains are computed differently when using the standard and mixed FE formulation. For this reason, when introducing the AFR technique, the consistency of the strain (and stress) fields computed with both formulations needs to be addressed.

As indicated in Fig. 3, in the standard displacement-based formulation, strains are evaluated at the integration points through local differentiation at element level using the nodal displacements

$$\epsilon \cong \hat{\epsilon} = \mathbf{S}\hat{\mathbf{u}} = \mathbf{S}\mathbf{u}_n \mathbf{U} = \mathbf{B}_u \mathbf{U} \quad (24)$$

Then, stresses are locally computed using these strains according to the constitutive law of the material.

Contrariwise, in the mixed formulation, strains are additional primary unknowns, calculated at the nodes. Then, the strains at the integration points are computed through the interpolation of the nodal values, as shown in Fig. 3, and stresses are also computed locally

$$\epsilon \cong \hat{\epsilon} = \mathbf{N}_e \mathbf{E} + \tau_e (\mathbf{B}_u \mathbf{U} - \mathbf{N}_e \mathbf{E}) = (1 - \tau_e) \mathbf{N}_e \mathbf{E} + \tau_e \mathbf{B}_u \mathbf{U} \quad (25)$$

3.1.2. Refinement criterion

Several refinement criteria may be adopted in the AFR strategy. The general requirements that these need to fulfill are:

1. To be adaptively applicable: the criterion should not require the a priori definition of the domains in which standard and mixed formulations are active. The areas of standard and mixed formulations should be automatically defined by the criterion.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

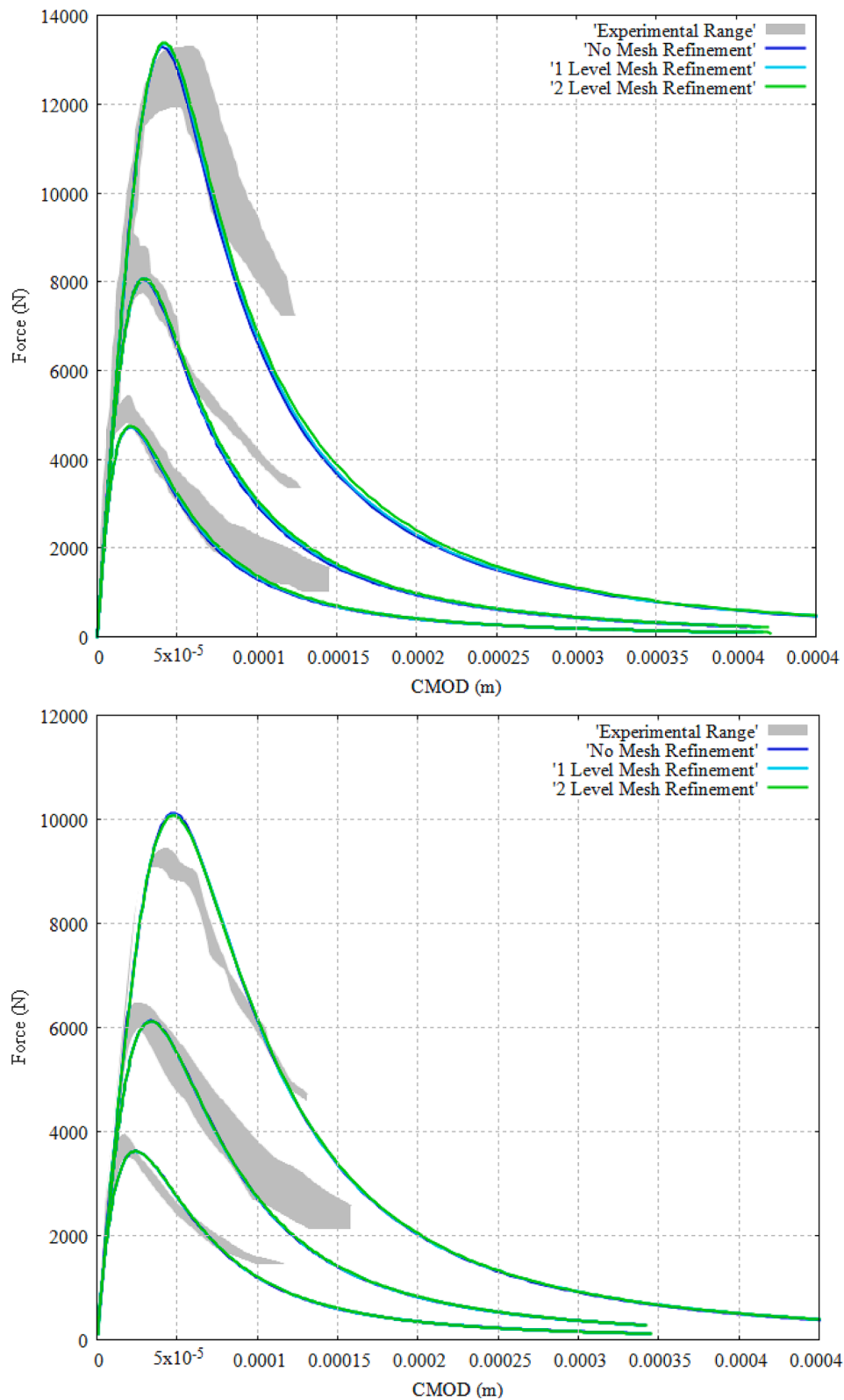


Fig. 9. Force-displacement curves of the Garcia-Alvarez beams for the 3 sizes using different levels of adaptive mesh refinement. Eccentricities (top) 0.625D and (bottom) 0.3125D.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

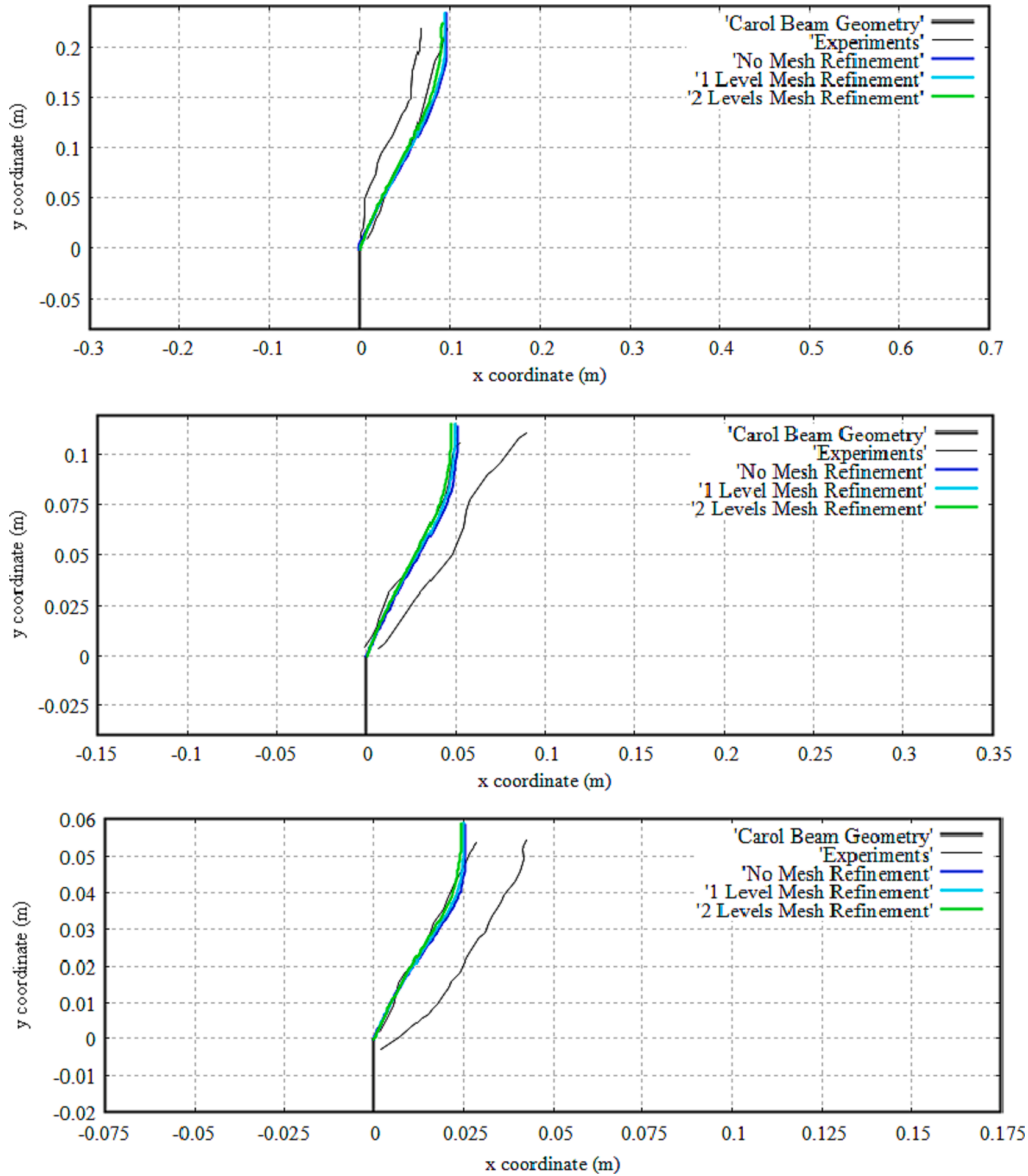


Fig. 10. Computed crack trajectories of the Garcia-Alvarez beams with different levels of mesh refinement, eccentricity 0.625D, for sizes (top) D = 320 mm, (center) D = 160 mm and (bottom) D = 80 mm.

2. To allow decision on both *formulation refinement*, switching from standard to mixed, and *formulation derefinement*, reversal from mixed to standard.
3. To designate the areas of standard and mixed formulations in an accurate and efficient manner: to reduce as much as possible the extent of the mixed formulation while preserving it in the required regions.

The criterion adopted to switch between the standard and mixed formulations is established in terms of the equivalent effective stress $\bar{\sigma}_{eq}$, introduced in the damage constitutive law in Section 2. In the present work the Rankine and Drucker-Prager criteria are adopted and $\bar{\sigma}_{eq}$ is determined according to Eqs. (19) and (20) respectively. Then, the ratio R is defined as

$$R = \frac{\bar{\sigma}_{eq}}{f_t} \quad (26)$$

where f_t is the tensile strength of the material.

If the ratio R exceeds a certain limit α_{max} (e.g. 75%), the mixed formulation is activated. Also, when R drops below a certain threshold α_{min} (e.g. 20 %), the standard formulation is recovered. The switching back from mixed to the standard formulation is not allowed in elements where damage has developed. Therefore, the criterion can be expressed as

IF $\bar{\sigma}_{eq} \geq \alpha_{max} f_t \rightarrow$ Activation of mixed FE

IF $\bar{\sigma}_{eq} < \alpha_{min} f_t$ AND $d = 0 \rightarrow$ Activation of standard FE (27)

This AFR criterion automatically defines the regions where the standard and mixed formulations are activated. If in an element the refinement criterion is fulfilled in *one* of the integration points, the whole element is treated as mixed and the strain degrees of freedom are activated in all its nodes. Conversely, if in *all* the integration points of an element the derefinement criterion is met, the standard formulation is

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

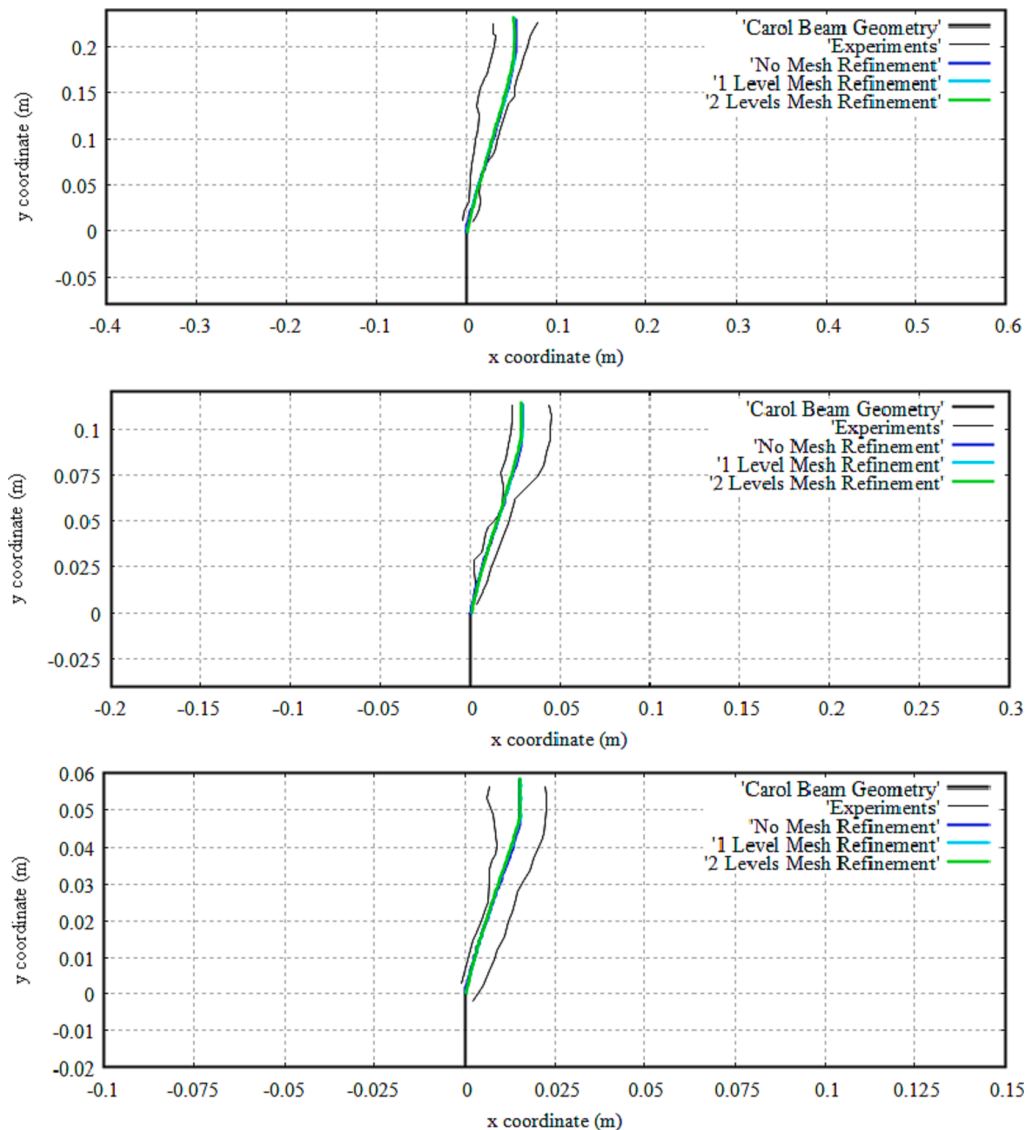


Fig. 11. Computed crack trajectories of the Garcia-Alvarez beams with different levels of mesh refinement, eccentricity 0.3125D, for sizes (top) $D = 320$ mm, (center) $D = 160$ mm and (bottom) $D = 80$ mm.

reactivated for that particular element. Note that the criterion allows the possibility of having several distinct subdomains with standard and mixed elements. Note also that in the original work [22] formulation derefinement was not contemplated.

3.1.3. Adaptive formulation solution approach

The approach employed to solve the problem with Adaptive Formulation Refinement is the following. According to the proposed refinement criterion, an array indicating for each node which degrees of freedom are active and inactive is updated at each time-step. Elemental matrices are computed and assembled in function of the active formulation. If in a certain node the mixed formulation is active, both displacement and strain degrees of freedom are considered. Instead, if the standard formulation is active, the strain degrees of freedom are not activated and only the displacement degrees of freedom are taken into consideration. In this strategy, all the nodes in the interface are treated as mixed, as can be seen in Fig. 2. The proposed methodology does not require additional considerations for interface nodes; contributions from different elements are simply assembled in the standard manner.

The solver stores the global matrix in a compressed format. In this way, only the non-zero entries of the matrix, which are located in function of its graph, are considered. The graph of the matrix is built by DOFs, according to the array labeling the active degrees of freedom of the problem. As demonstrated in reference [22], this approach greatly reduces the computational time involved, as only the active degrees of freedom are considered when building and solving the problem. This method is applicable both for direct and iterative solvers.

3.2. Adaptive mesh refinement

The AMR strategy is included to further increase the efficiency of the model for fracture conferred by the AFR approach. The AMR method allows to begin the computation with an initial relatively coarse mesh requiring a low calculation effort. The mesh is progressively and adaptively refined only in the regions where cracks develop to suitably represent the phenomenon of fracture while the coarse mesh is maintained elsewhere. The AMR methodology also includes the subsequent adaptive mesh coarsening of the undamaged elements after the crack

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

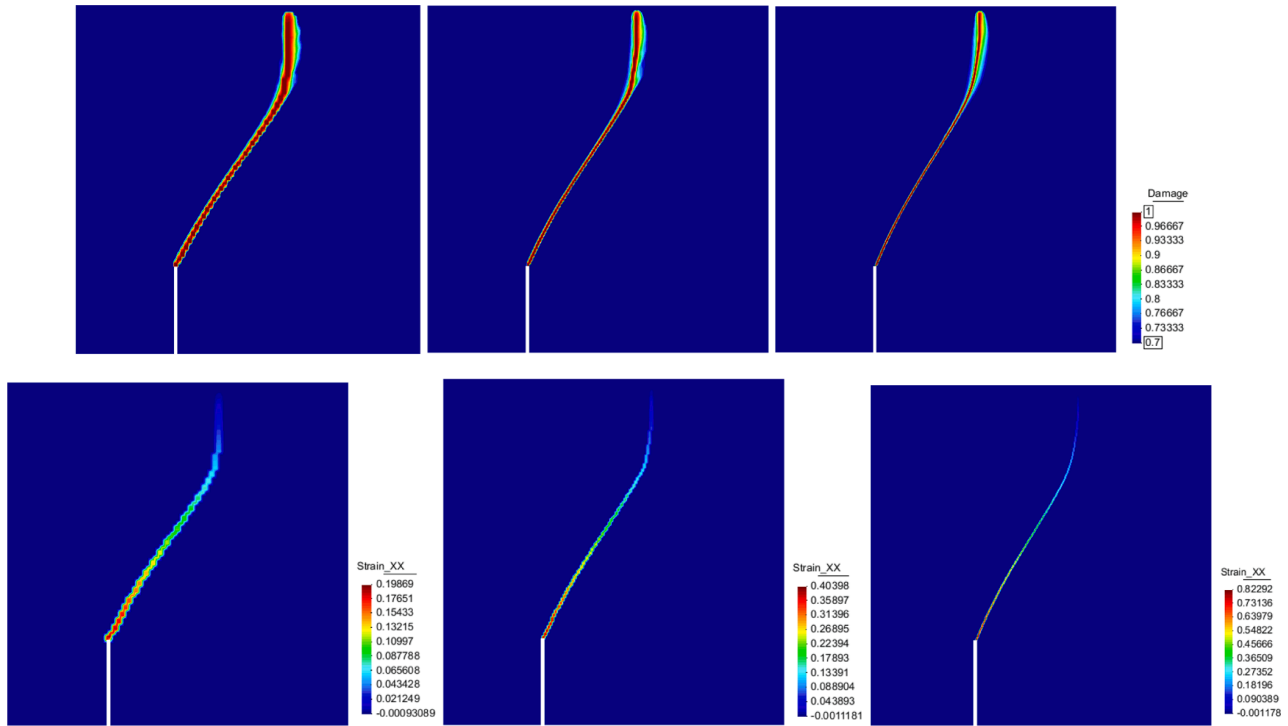


Fig. 12. Computed (top) damage and (bottom) strain contours of the Garcia-Alvarez beams, $D = 80$ mm specimen, eccentricity 0.625D series, with (left) no AMR, (center) 1 level of AMR and (right) 2 levels of AMR.

has already appeared in the area. An octree refinement strategy is employed, according to the methodology described in references [28,29,30,31]. In problems where the crack trajectories are not known a priori this approach avoids the need for initial fine meshes to calculate fracture.

3.2.1. Refinement criterion

A similar refinement criterion as in AFR, based on the same ratio R between the equivalent effective stress $\bar{\sigma}_{eq}$ and the material tensile strength f_t , as defined by Eq. (26), is adopted for AMR. If the ratio R exceeds a certain designated limit β_{max} (e.g. 50%), the mesh is refined. Also, if the ratio R decreases below another set value β_{min} (e.g. 10%), coarsening takes place. Refinement may take place up to a certain pre-determined level. Coarsening is not allowed in damaged elements. Hence, the criterion can be stated in the following way:

IF $\bar{\sigma}_{eq} \geq \beta_{max} f_t \rightarrow$ Mesh refinement

IF $\bar{\sigma}_{eq} < \beta_{min} f_t$ AND $d = 0 \rightarrow$ Mesh coarsening (28)

Note that the values of the thresholds introduced for the refinement criteria in AMR and AFR in the numerical simulations may be different. This means that AMR and AFR may be activated independently, depending on the selected α_{min} , α_{max} , β_{min} , β_{max} parameters.

Analogously to AFR, mesh refinement takes place in an element if the refinement criterion is reached in *one* of the integration points; an additional refinement level is introduced at every time step until the predefined maximum number of refinement levels is attained. Contrarily, coarsening occurs if *all* the integration points of the element meet the coarsening criterion.

3.2.2. Adaptive finite element approach

In the following, a brief overview of the FE adaptivity strategy employed is provided. For additional details, references [28,29,30,31] are recommended.

The hierarchical octree refinement strategy defines several levels of FEs, represented in Fig. 4 and Fig. 5. The original FEs are labeled with the level zero and are the parents to the level one FEs (which are the children of the zero level FEs). Level one FEs are the parents of level two FEs, etc.

When a new mesh is generated during the AMR process, the information from the previous mesh is projected onto the new one. The nodal displacements of existing nodes are preserved. The values corresponding to new nodes are interpolated from the values of existing nodes. If the mixed FE formulation is active for the parent element, the same procedure is followed for the nodal strain field. The historical variables of the damage constitutive law of new nodes are set to the values corresponding to the initially undamaged material. Since coarsening is allowed only in undamaged elements, no loss of information related to the historical variables of the nonlinear constitutive behavior occurs. Setting proper limits for the refinement criterion also avoids the refinement of already damaged elements. With regard to the combination of AMR with AFR, the array indicating which degrees of freedom are active and inactive for each node is then updated.

The presence of hanging nodes in the octree refined FE mesh, depicted in Fig. 6, needs to be properly treated. In the present work, the treatment of hanging nodes proposed in references [28,29] is followed. According to this approach, the following modification of the shape functions of the parent elements is performed when hanging nodes are present so that the FE approximation becomes conforming:

$$N_p^*(x) = N_p(x) + \sum_{i=1}^{n_{child}} N_p(x_i) N_{c,i}(x) \quad (29)$$

where N_p and N_p^* are the original and modified parent interpolation functions, N_c are the original children interpolation function, n_{child} is the number of hanging children nodes and x_i indicates the position of hanging child i . The degrees of freedom of hanging nodes are eliminated from the FE system being solved. Their contributions are added to the nodes of the parent elements according to the modified shape functions,

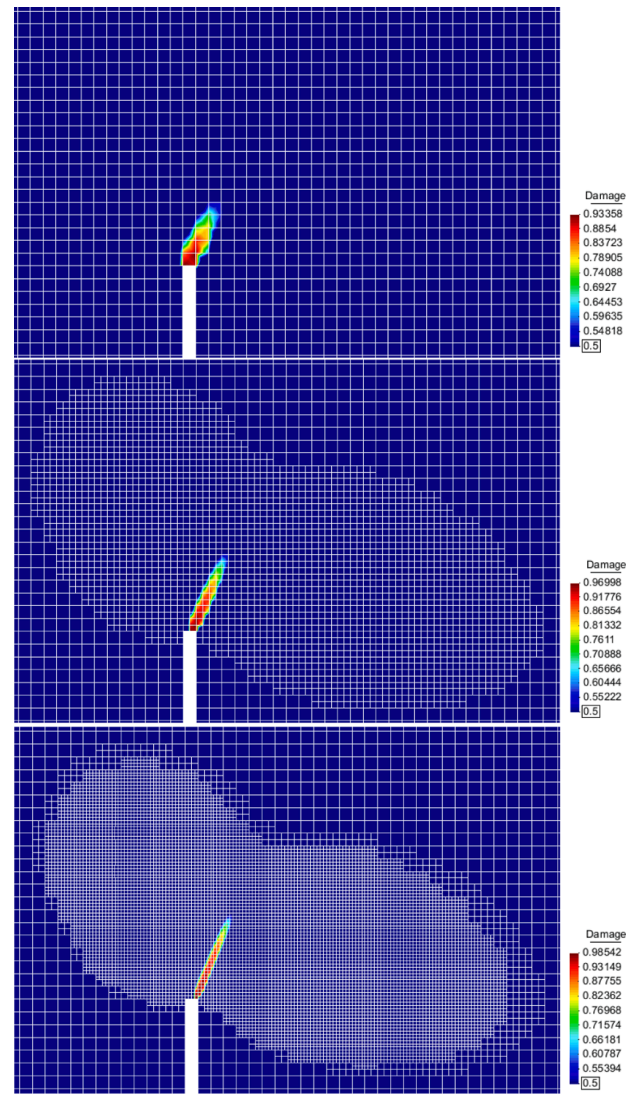


Fig. 13. Computed damage contour with the detail of the mesh in the notched region at the early stages of the simulation with (top) no AMR, (center) 1 level of AMR and (bottom) 2 levels of AMR.

Eq. (29). The values of the unknowns at the hanging nodes are obtained by interpolation from nodes of the parent element. Additional details are provided in the original references [28,29].

The combination of the AFR and AMR methodologies is fairly straightforward and does not require special additional considerations.

Table 2
Cost-efficiency of the proposed mesh and formulation adaptivity scheme for the Garcia-Alvarez beams.

	Max. Ndofs standard FEs	Max. Ndofs mixed ϵ/u FEs	Max. Ndofs AFR	% of Ndofs with respect to mixed only	% of Ndofs with respect to standard only
Without AMR	63,226	189,678	71,222	37.55%	112.65%
1 level of AMR	89,196	267,588	100,012	37.38%	112.13%
2 levels of AMR	190,926	572,778	229,870	40.13%	120.40%

Table 3
Cost-efficiency of the proposed mesh adaptivity scheme for the Garcia-Alvarez beams.

	Number of nodes with initial uniform refinement	Max. number of nodes with AMR	%
No mesh refinement	31,613	–	–
1 level of refinement	125,575	44,598	35.52%
2 levels of refinement	500,549	95,463	19.07%

4. Numerical simulation of quasi-brittle structural failure

In this section the numerical simulation of experimental tests and benchmark case studies reported in the literature featuring quasi-brittle structural failure is performed using the proposed model including mesh and FE formulation adaptivity. The selected cases are often employed for the validation of numerical models.

The objectives of this section are (i) to demonstrate the mesh objectivity of the proposed model based on the AFR and AMR approaches, (ii) to assess the cost-efficiency of the proposed method compared to the standard displacement-based FEs as well as the non-adaptive mixed FEs, (iii) to show proper convergence of the results to the appropriate solution.

The computations are performed with an enhanced version of the finite element code FEMUSS while pre- and post- processing are performed with GID [52], both developed at the International Center for Numerical Methods in Engineering (CIMNE). Convergence in each load step is attained when the ratio between the norm of residual forces and the norm of total external forces is lower than 10^{-5} .

4.1. Garcia-Alvarez beams

In this section, the numerical simulation of the Garcia-Alvarez experiments, reported in reference [53], is performed. The work [53] includes the computational modelling of the tests using interface elements. This problem has also been simulated with the phase-field

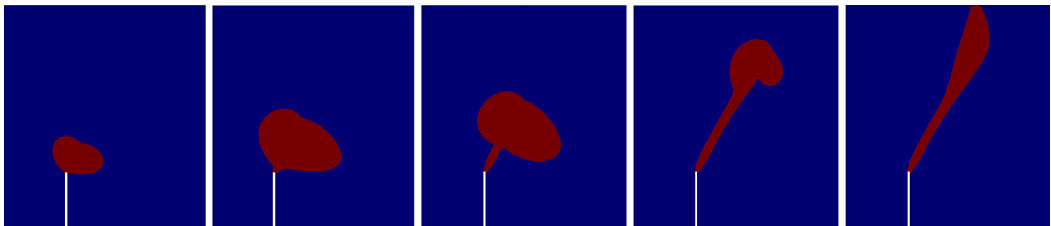


Fig. 14. Evolution of the area where standard (in blue) and mixed (in red) formulations are adopted during the simulation with 2 levels of AMR of the Garcia-Alvarez beams, eccentricity 0.625D, D = 80 mm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

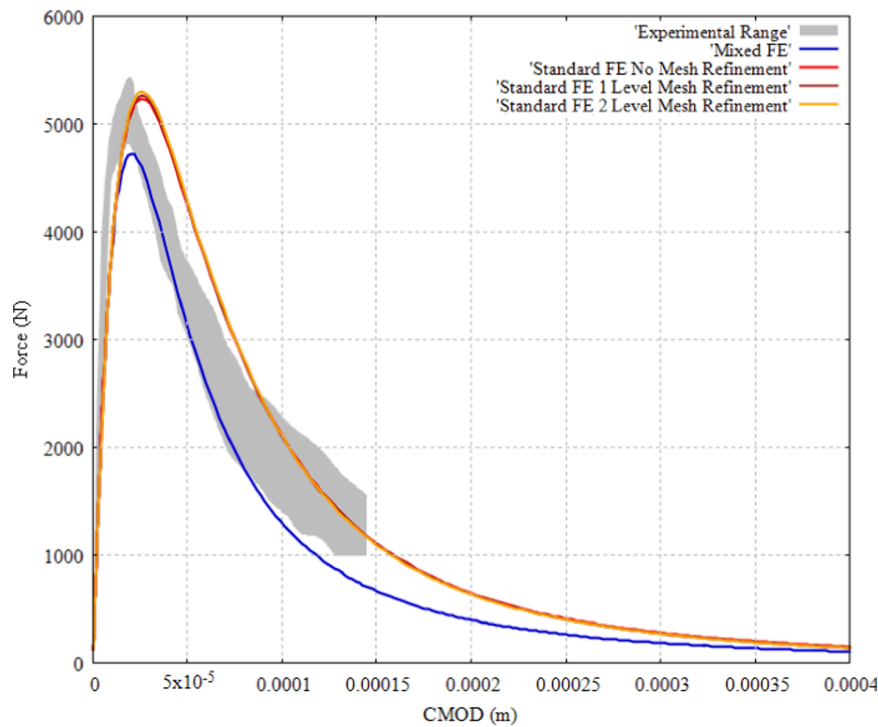


Fig. 15. Force-displacement curves of the Garcia-Alvarez beams, standard FEs, using different levels of adaptive mesh refinement. $D = 80$ mm specimen and eccentricity $0.625D$.

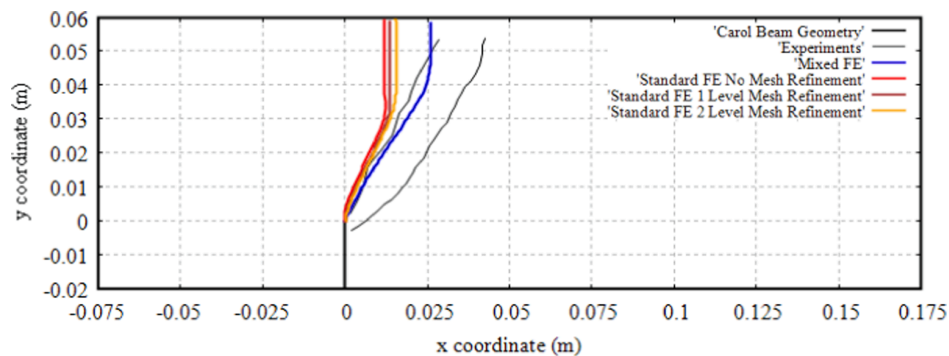


Fig. 16. Computed crack trajectories of the Garcia-Alvarez beams, standard FEs, using different levels of adaptive mesh refinement. $D = 80$ mm specimen and eccentricity $0.625D$.

regularized cohesive zone model (PF-CZM) by Wu and coworkers [54]. In previous works the results obtained using mixed FE have been compared with the experiments [15] and with phase-field based computations [6,55].

In these experiments, geometrically similar beams of different sizes are subjected to three-point bending tests in order to study the structural size effect. The geometry of the beams is detailed in Fig. 7. Three different structural sizes are considered: beams with depths D of 320 mm, 160 mm and 80 mm. A constant span-to-depth ratio of 2.5 is taken. In the experiments three different series are considered, by varying the position of the notch in the specimen, so that mixed-mode fracture occurs. Specifically, notch eccentricities μD of $0.625D$, $0.3125D$ and $0.0D$ were used. The thickness of the beams is 50 mm and the notch-to-depth ratio λ is 0.25 in all the cases. A vertical force was applied at the top midpoint of the beams as indicated in Fig. 7. The computational simulations are performed under CMOD control, reproducing the conditions of the experiment. Table 1 shows the material properties used for the

beams, together with the Rankine damage surface.

The objective of this section is to validate the proposed AMR and AFR capabilities introduced in this work. For this, the numerical simulation of the series with notch eccentricities μD of $0.625D$ and $0.3125D$ is performed, with concurrent adaptive FE formulation refinement and different levels of adaptive mesh refinement, up to 2 levels.

The initial meshes adopted to perform the simulation, for the two eccentricity cases and all the sizes, have approximately 31,000 elements, resulting from using a mesh size h of $10^{-2}D$. The notch is modelled with thickness h , as is shown in Fig. 8. The original mesh is fine enough to obtain satisfactory results; it is chosen to show that proper convergence occurs with different levels of adaptive mesh refinement.

In Fig. 9, the force-displacement curves computed for the two series and for all the sizes are shown. It can be seen that the results with the different levels of adaptive mesh refinement (from no mesh refinement to 2 levels of mesh refinement) are practically overlapping. This shows that results have reached convergence in terms of mesh size. All results

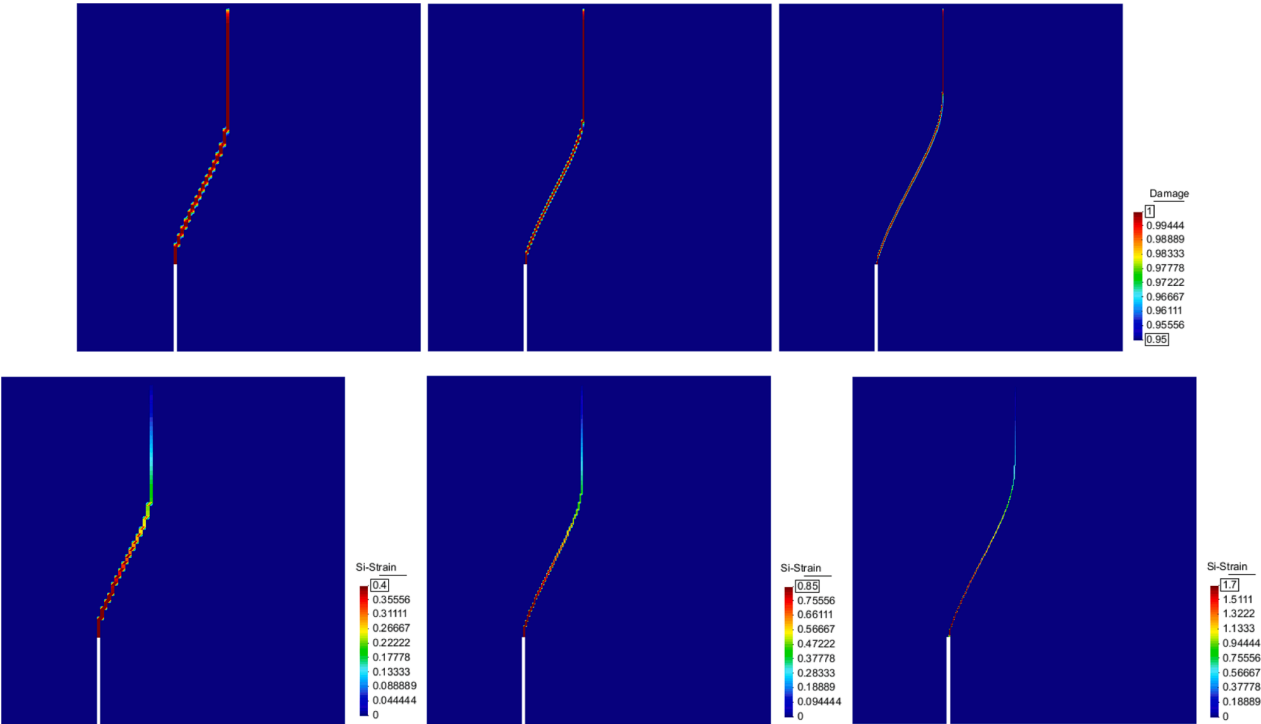


Fig. 17. Computed (top) damage and (bottom) maximum principal strain contours using standard FEs, $D = 80$ mm specimen, eccentricity $0.625D$ series, with (left) no AMR, (center) 1 level of AMR and (right) 2 levels of AMR.

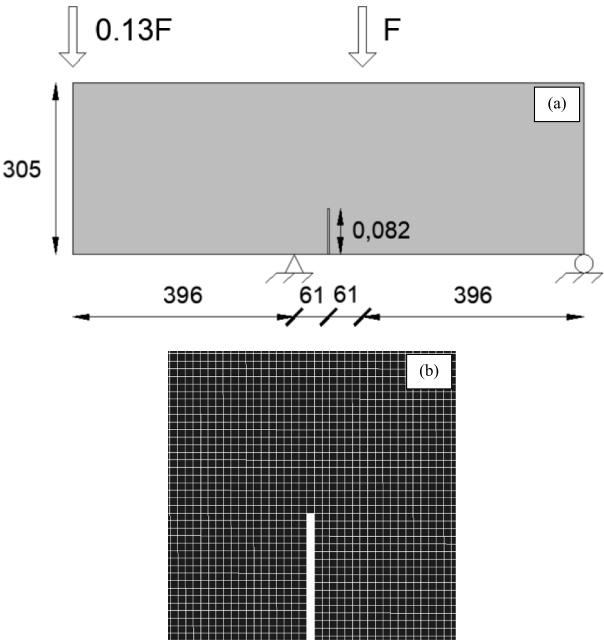


Fig. 18. Arrea and Ingrassia beam: (a) Geometry (mm) and (b) detail of the mesh in the area of the notch.

compare well with the experiments for all the sizes and for all the cases. This is remarkable, as for all the simulations the same material properties are used. This confirms the results in reference [15] where the ability of the mixed finite element method in reproducing the structural size effect is demonstrated.

Table 4
Material parameters of the Arrea and Ingrassia beam.

Young's Modulus	$28.8 \cdot 10^9$ Pa
Poisson's Ratio	0.18
Tensile Strength	$3.7 \cdot 10^6$ Pa
Tensile Fracture Energy	130 J/m ²
Compressive Strength	$43.4 \cdot 10^6$ Pa

Fig. 10 and Fig. 11 compares the computed crack trajectories for all the sizes of the two beam series with the corresponding experiments. It can be seen that all the numerical results are within the experimental range. The simulations with the different levels of mesh refinement considered are practically matching, demonstrating that convergence of the results has been achieved. No sign of spurious mesh dependence is observed when using the AFR method combining standard and mixed FEs.

In Fig. 12 the computed damage and strain fields for the $D = 80$ mm sized specimen, eccentricity $0.625D$, are displayed. A comparison between results obtained with (i) no AMR, (ii) 1 level of AMR and (iii) 2 level of AMR is made. In all cases, damage is plotted within the range that matches the corresponding localizing strains. It can be seen both in the damage and strain fields how the crack bandwidth is reduced when the FE size decreases. Notwithstanding, the correct energy dissipation with respect to the FE size is enforced by the Crack Band Theory, as it can be seen in the force–displacement curves in Fig. 9; also, practically the same crack trajectories are obtained, as shown in Fig. 10.

Fig. 13 shows the detail of the FE mesh in the region of the notch at the early stages of the simulation without AMR and with 1 and 2 levels of AMR. It can be seen how the FE mesh is refined only in the area where the crack progresses according to the adaptive mesh refinement criterion defined in Section 3.

In the computations of the Garcia-Alvarez beams, mesh refinement takes place when the equivalent effective stress $\bar{\sigma}_{eq}$ reaches a value higher than $0.5f_t$ and coarsening occurs when $\bar{\sigma}_{eq}$ attains a value lower

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

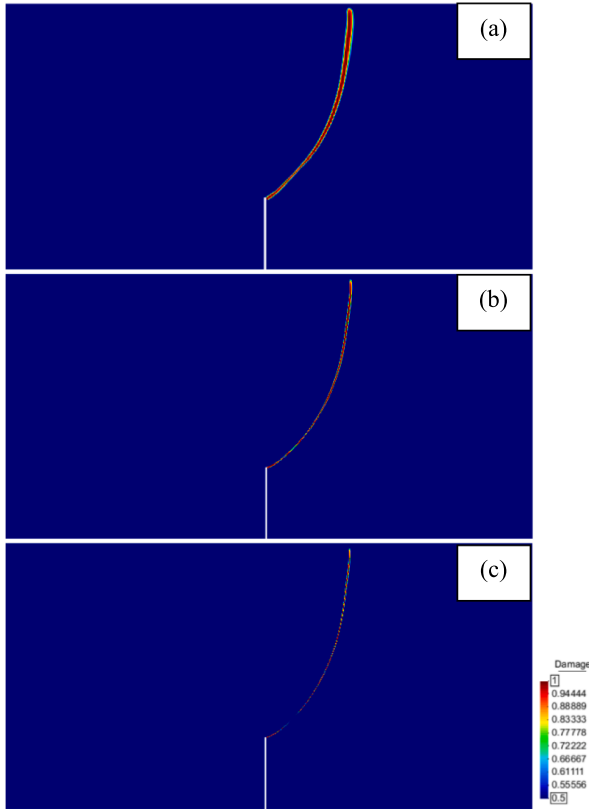


Fig. 19. Computed damage contours of the Arrea and Ingraffea beam with: (a) no AMR, (b) 1 level of AMR and (c) 2 levels of AMR.

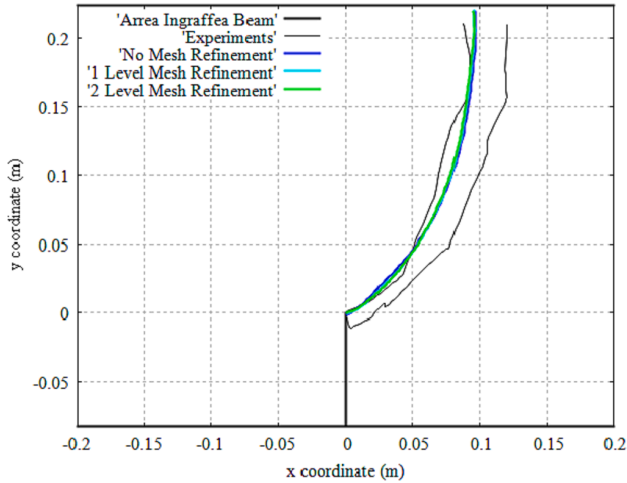


Fig. 20. Detail of the computed crack trajectories for the different levels of mesh refinement of the Arrea and Ingraffea beam.

than $0.1f_t$.

Fig. 14 shows the evolution of the area where standard and mixed FE formulations are active for different time steps of the simulation of the Garcia-Alvarez beam, $D = 80 \text{ mm}$ sized specimen, eccentricity $0.625D$. With the criterion defined in Section 3, mixed ϵ/u FEs are activated only in a small region near the developing crack.

In all the simulations of the Garcia-Alvarez beams, formulation

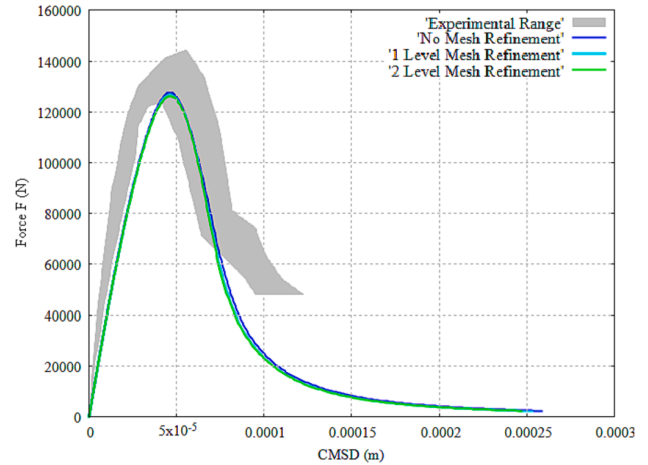


Fig. 21. Computed Force F vs CMSD curves of the Arrea and Ingraffea beam.

refinement, i.e. activation of the mixed ϵ/u FE, occurs when the equivalent effective stress $\bar{\sigma}_{eq}$ reaches a value higher than $0.75f_t$. Formulation derefinement, i.e. deactivation of the mixed ϵ/u FE, takes place when $\bar{\sigma}_{eq}$ attains a value lower than $0.2f_t$; no switching from mixed to standard takes place in damaged elements.

In Table 2 the details of the reduction in computational cost by the adaptive formulation refinement methodology are indicated for the $D = 80 \text{ mm}$ and eccentricity $0.625D$ case. Similar values are obtained for all the other cases of the Garcia-Alvarez beams. For each of the cases considered (without AMR, with 1 level of AMR and with 2 levels of AMR) the first three columns compare the maximum number of degrees of freedom (labelled Max. Ndofs) ensuing from computations with standard FE only, mixed FE only, and the proposed AFR strategy. The last two columns show the relation between the Max. Ndofs resulting from the use of AFR with respect to the corresponding values of standard FE only and mixed FE only computations.

In 2D and 3D simulations, the ratio of unknowns between the standard and mixed FE formulations is 33.33%. This means that the lowest possible value that can be reached in the fourth column of Table 2 for the percentage of Max. Ndofs using AFR with respect the a mixed only calculation is 33.33%. Note that the actual values indicated in the column are very close to this lower bound, demonstrating the efficiency of the proposed approach. The fifth column of Table 2 indicates the relation between the Max. Ndofs using AFR with respect the standard only analysis, which is very close in all the cases to the lowest possible value of 100%. The use of the AFR allows to decisively reduce the number of degrees of freedom, and therefore the computational cost, of the simulation with respect to the use of the mixed ϵ/u formulation. In fact, the maximum number of degrees of freedom required with AFR is only slightly larger than that of using the standard FE formulation.

Table 3 further details the cost-effectiveness of the proposed mesh adaptivity scheme for the numerical simulation of quasi-brittle fracture. For each of the cases considered (without refinement, with 1 level of refinement and with 2 levels of refinement) the first column indicates the number of nodes that ensue from performing an initial uniform refinement of the mesh and the second column displays the number of nodes resulting from the adoption of the AMR strategy. The third column presents the relation between the two previous values. It can be seen how the adopted AMR scheme substantially reduces the number of nodes, and therefore the computational cost, of the simulation with respect to the alternative of performing an initial uniform refinement of the mesh.

The calculations show that: (i) a large reduction in computational cost by the use of the adaptive FE formulation capabilities is achieved, (ii) results are convergent when different levels of adaptive mesh

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

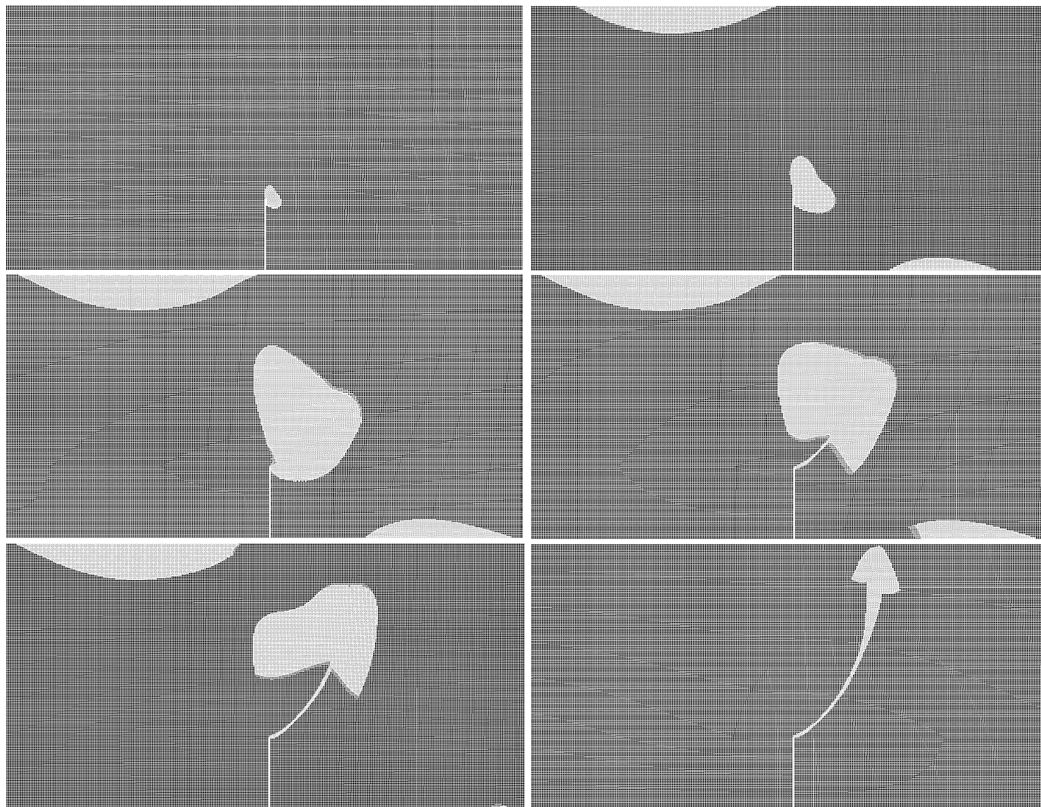


Fig. 22. Evolution of the FE mesh during the simulation with 2 levels of AMR of the Arrea and Ingraffea beam.

refinement are considered, both in terms of crack trajectories and force–displacement curves, (iii) results agree with the experiments and do not exhibit any sign of spurious mesh dependency, (iv) the quasi-brittle fracture model based on the adoption of the mixed strain/displacement FE method is able of reproducing structural size effect with accuracy.

4.1.1. Comparison with standard FE simulations

In the following, results obtained using standard FEs only are compared with the ones of the previous section. The objective is to demonstrate that mesh refinement by itself is not able to yield satisfactory results; the use of mixed ϵ/u FEs is necessary for that purpose. In this section the specimen with $D = 80 \text{ mm}$ and eccentricity $0.625D$ is considered only. The same initial mesh and material properties are employed.

Fig. 15 shows the force–displacement curves computed with standard FEs using different levels of AMR. It can be observed that, despite the fact that the solutions computed with standard FEs are practically overlapping, they converge to a different force–displacement curve than mixed FEs.

In Fig. 16, the crack trajectories computed with standard FEs are shown. It can be observed that these vary when different levels of mesh refinement are adopted, showing spurious mesh bias dependency and non-convergent results. Also, fractures calculated with standard FEs are outside the experimental range. The large difference between the crack trajectories obtained with standard and mixed FEs is the reason for obtaining different solutions in terms of force–displacement curves as well, the difference between the two calculated peak loads being of 11%.

Fig. 17 displays the computed damage and maximum principal strain fields using standard FEs, comparing the results without AMR and with 1 and 2 levels of AMR. Again, it can be seen how the crack bandwidth diminishes as the FE size decreases. Despite the fact that the resulting

force–displacement curves are converging to the same solution when using standard FEs, showing that the correct energy dissipation with respect to the FE size is enforced by the Crack Band Theory, the computed results are different in terms of crack trajectory, as shown in Fig. 16.

4.2. Arrea and Ingraffea beam

This section presents the numerical simulation of the Arrea and Ingraffea mixed mode bending test, reported in reference [56]. As reference [6] details, this experiment has been numerically modelled many times using several different approaches to crack modelling.

The geometry and the boundary conditions of the beam are indicated in Fig. 18a. The eccentricity of the applied loading with respect to the notch results in the mixed mode of the fracture. The beam has dimensions of $0.914 \text{ m} \times 0.305 \text{ m}$ and a notch length of 82.4 mm . The beam thickness is 0.152 m . Two eccentric loads of values F and $0.13F$ act on the upper surface. The material properties considered are indicated in Table 4. The Drucker-Prager damage surface is considered for the damage constitutive law in this case. The numerical simulations are performed under Crack Mouth Sliding Displacement (CMSD) control.

The objective of this section is to validate the proposed methodology through comparison with the experimental results. To this end, the present case study examines the results of the numerical simulations obtained with (i) no mesh refinement, (ii) 1 level of adaptive mesh refinement and (iii) 2 levels of adaptive mesh refinement. AMR takes place when the equivalent effective stress $\bar{\sigma}_{eq} \geq 0.5f_t$. The coarsening of the mesh is decided when $\bar{\sigma}_{eq} \leq 0.1f_t$. In all the cases, the adaptive formulation refinement capabilities are introduced as well. The mixed formulation is activated if $\bar{\sigma}_{eq} \geq 0.75f_t$ and the standard formulation is reestablished if $\bar{\sigma}_{eq} \leq 0.5f_t$.

The initial mesh employed has a FE size of 2 mm . This corresponds to

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

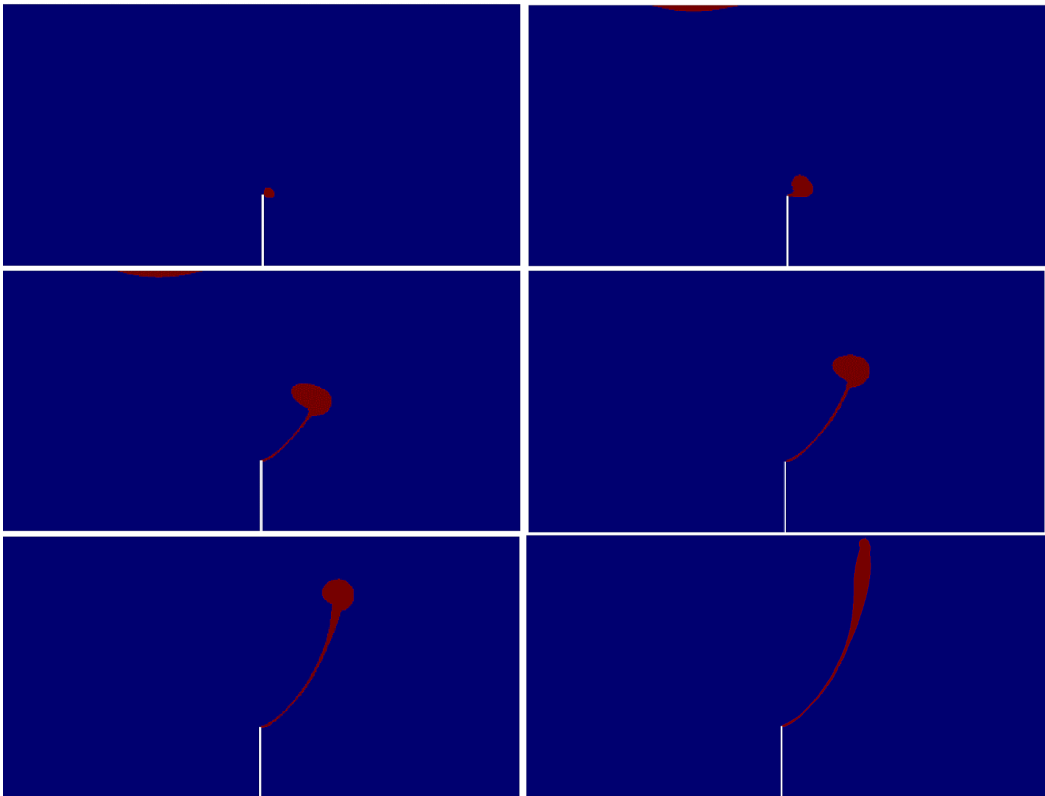


Fig. 23. Evolution of the area where standard (in blue) and mixed (in red) formulations are adopted during the simulation with 2 levels of AMR of the Arrea and Ingraffea beam. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Cost-efficiency of the proposed mesh and formulation adaptivity scheme for the Arrea and Ingraffea beam.

	Max. Ndofs standard FEs	Max. Ndofs mixed ϵ/u FEs	Max. Ndofs AFR	% of Ndofs with respect to mixed only	% of Ndofs with respect to standard only
Without AMR	144,910	434,730	153,050	35.21%	105.62%
1 level of AMR	182,566	547,698	192,786	35.20%	105.60%
2 levels of AMR	329,890	989,670	358,350	36.21%	108.63%

Table 6
Cost-efficiency of the proposed mesh adaptivity scheme for the Arrea and Ingraffea beam.

	Number of elements with initial uniform refinement	Max. number of elements with AMR	%
No mesh refinement	71,761	–	–
1 level of refinement	287,044	90,724	31.61%
2 levels of refinement	1,148,176	164,077	14.29%

a mesh of 71,761 elements. The notch is modelled with a thickness of 1 finite element, as is shown in Fig. 18b. The initial mesh considered in the simulations is fine enough to provide satisfactory results; it is selected to demonstrate that convergence occurs when adaptive mesh refinement is activated.

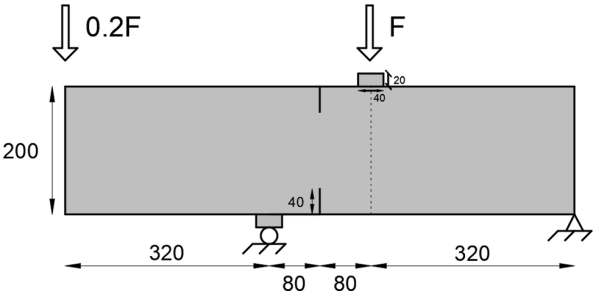


Fig. 24. Geometry of the Bocca beam (mm).

Table 7
Material properties of the Bocca beam.

Young's Modulus	27.0·10 ⁹ Pa
Poisson's Ratio	0.18
Tensile Strength	2.0·10 ⁶ Pa
Tensile Fracture Energy	100 J/m ²
Compressive Strength	37.3·10 ⁶ Pa

Fig. 19 shows the computed damage contours, comparing results (i) without AMR, (ii) with 1 level of AMR and (iii) with 2 levels of AMR. The crack, starting at the notch, gradually turns and modifies its trajectory due to the mixed mode loading conditions of the test. The detail of the crack trajectories is examined in Fig. 20. On the one hand, it can be seen how the numerical results are practically overlapping and that convergence of the solution is properly taking place upon mesh refinement. On the other hand, the crack trajectories obtained with the adaptive FE

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

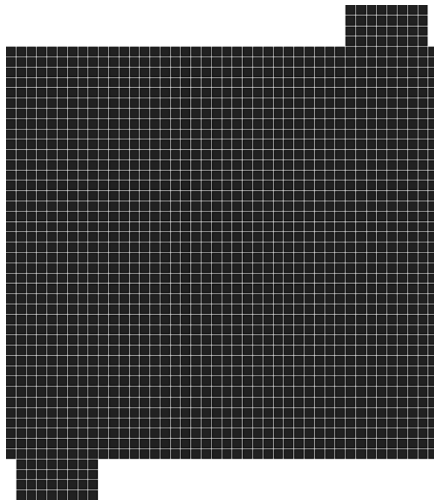


Fig. 25. Detail of the initial mesh employed in the analysis of the Bocca beam in the central area.

formulation method proposed in this work matches the experimental observations.

In Fig. 21 the computed Force F vs CMSD curves with the different levels of mesh refinement are displayed. It can be seen again that proper convergence of the solution is taking place as well when considering the force–displacement response. The computations agree with the experimental observations both in terms of bearing capacity and of overall nonlinear behavior.

Fig. 22 shows the evolution of the FE mesh during the simulation as a result of the adaptive mesh refinement capabilities introduced in the model. In Fig. 23 the evolution of the areas where standard and mixed

FEs are active during the simulation is presented. The ability of the criterion based on $\bar{\sigma}_{eq}$ to allow the correct switching between standard and mixed formulations effectively is verified again.

In Table 5 the cost-efficiency of the proposed AFR scheme is assessed. For the three cases considered (without AMR, with 1 level of AMR and with 2 levels of AMR), the first three columns provide the maximum number of degrees of freedom (Max. Ndots) involved in the computations using standard FE only, mixed FE only and AFR respectively. The last two columns indicate the relation between the Max. Ndots in AFR with respect to the other two approaches, (standard only and mixed only). The analyzes performed with AFR necessitated a very small increment in the number of degrees of freedom with respect to the use of standard FEs (5%).

Table 6 presents the cost-effectiveness of the AMR approach. For each of the contemplated cases (without refinement, with 1 level of refinement and with 2 levels of refinement) the first column shows the number of elements resulting from the adoption of a uniform mesh refinement strategy. The second column displays the number of nodes that ensue from the use of the AMR method proposed in this work. The third column details the ratio between the two previous values. It can be observed how the proposed AMR strategy allows for a very significant reduction in the number of elements and therefore the cost of the computational simulations.

Computations with the mixed ϵ/u FE approach do not present any trace of spurious mesh dependency. The adaptive formulation approach allows to obtain results with enhanced accuracy and mesh objectivity while introducing an extremely reduced increase in the computational cost with respect to the use of standard FEs. The efficiencies displayed by the AFR and AMR methods are caused by the fact that the crack propagates in a reduced area of the whole structure. Only a small portion of the beam reaches the nonlinear regime, as it can be observed in Fig. 19, Fig. 22 and Fig. 23.

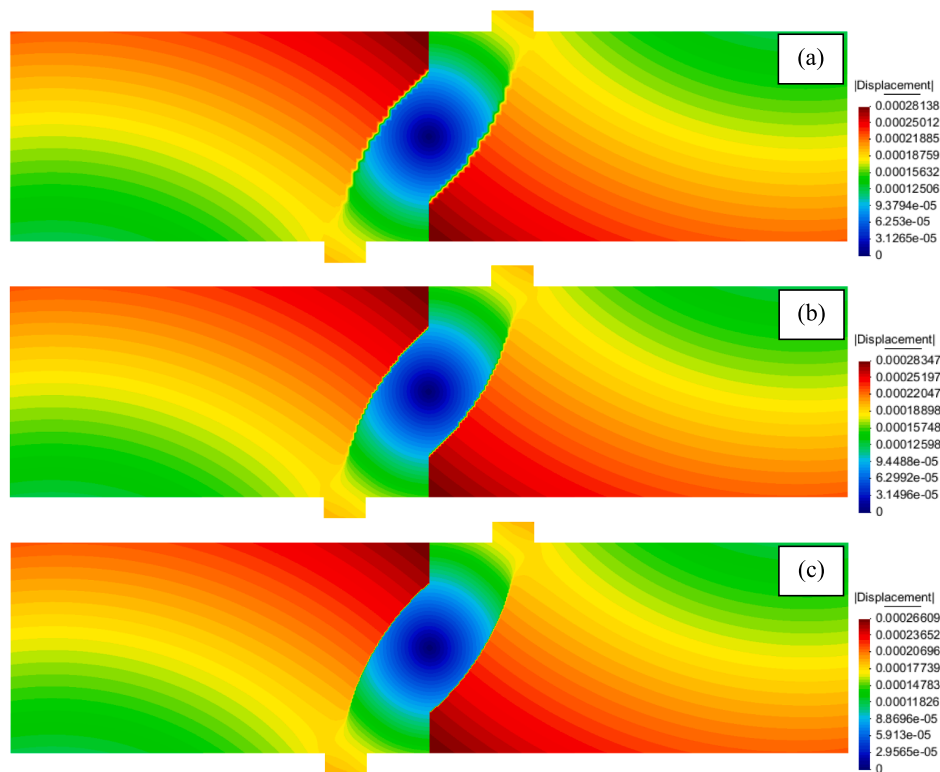


Fig. 26. Displacement contours of the Bocca beam with: (a) no AMR, (b) 1 level of AMR and (c) 2 levels of AMR.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

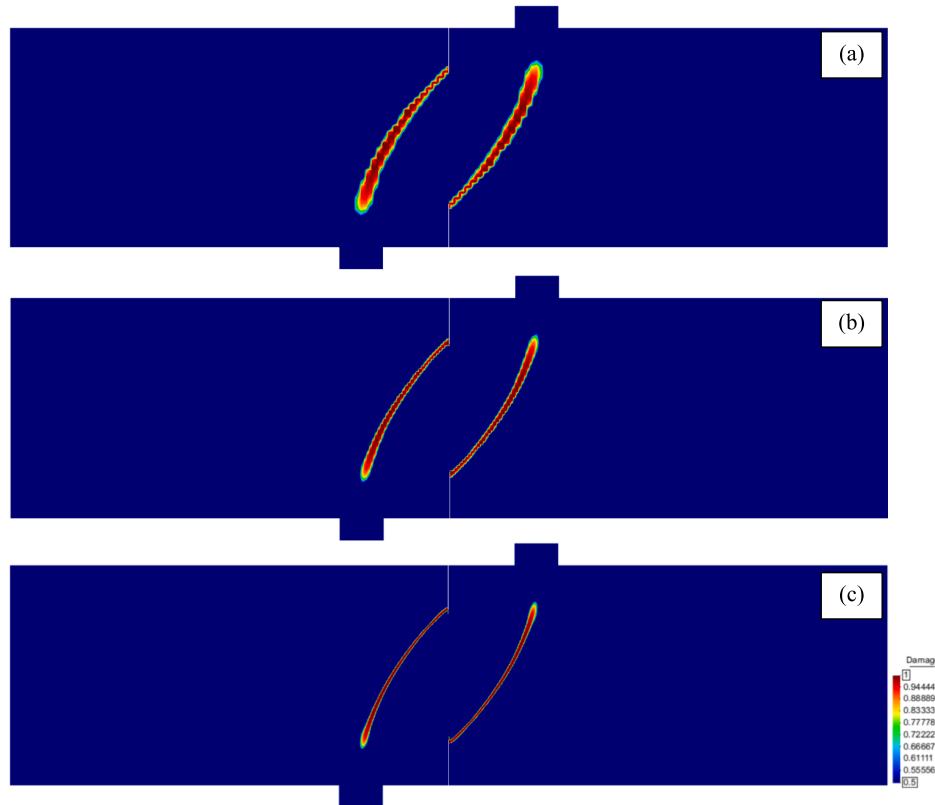


Fig. 27. Damage contours of the Bocca beam with: (a) no AMR, (b) 1 level of AMR and (c) 2 levels of AMR.

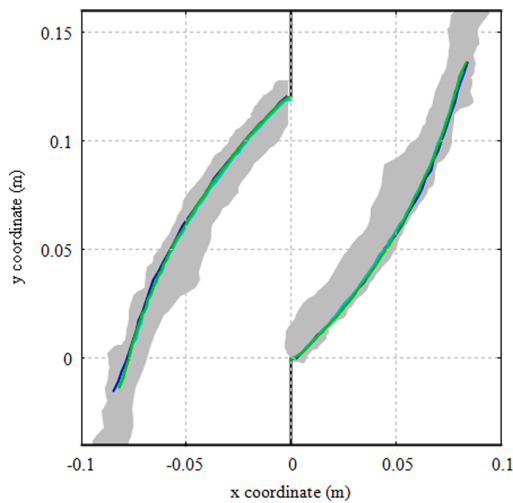


Fig. 28. Detail of the computed crack trajectories for the different levels of mesh refinement of the Bocca beam.

4.3. Bocca beam

This section analyzes the capabilities of the proposed model in reproducing the results of the doubly notched beam of the Bocca experiment, reported in reference [57]. Reference [57] includes as well the numerical simulation of the test using an automatic remeshing scheme to reproduce crack propagation.

This problem has been solved in references [58,59], where particle meshless methods are adopted, in reference [60], where the boundary

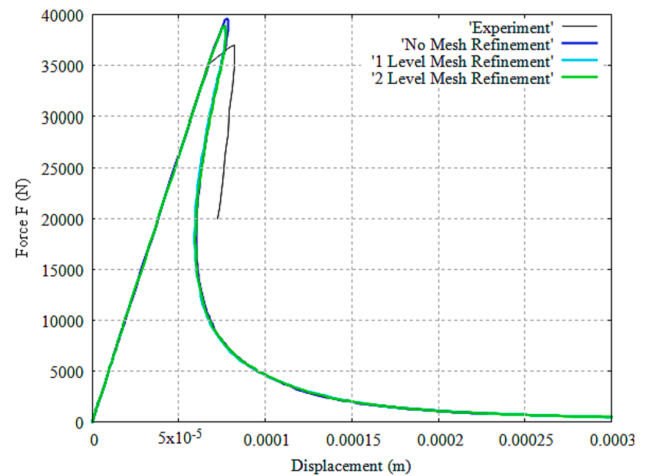


Fig. 29. Computed force-displacement curves of the Bocca beam.

element method is used, in reference [61], which employed EFEM, in reference [62], where another automatic remeshing scheme is proposed and in reference [63], which introduced a localization limiter to regularize the problem.

The geometry and boundary conditions of the test are detailed in Fig. 24. The beam has dimensions $0.8 \text{ m} \times 0.2 \text{ m}$ and a thickness of 0.1 m . Two notches of 40 mm length are introduced, and two eccentric forces of values F and $0.2F$ are applied as detailed in Fig. 24. Note that the beam geometry and boundary conditions are polar symmetric. The material parameters are given in Table 7. The Drucker-Prager damage criterion is adopted. The simulation is performed under crack mouth

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

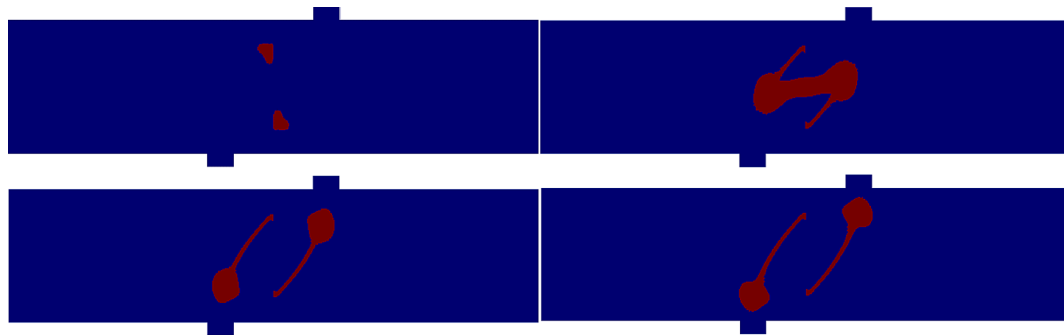


Fig. 30. Evolution of the area where standard (in blue) and mixed (in red) formulations are adopted during the simulation with 2 level adaptive mesh refinement of the Bocca beam. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

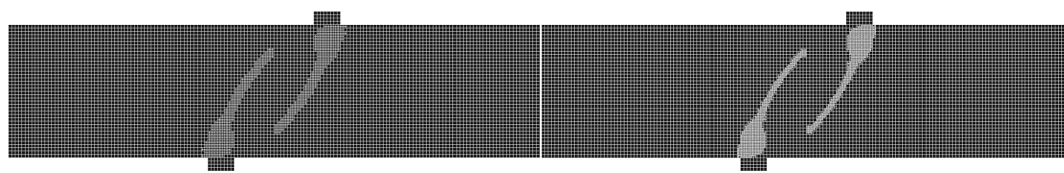


Fig. 31. Meshes at the end of the simulations of the Bocca beam with (left) 1 level of AMR and (right) 2 levels of AMR.

Table 8

Cost-efficiency of the proposed mesh and formulation adaptivity scheme for the Bocca beam.

	Max. Ndofs standard FEs	Max. Ndofs mixed ϵ/u FEs	Max. Ndofs AFR	% of Ndofs with respect to mixed only	% of Ndofs with respect to standard only
Without AMR	13,378	40,134	17,230	42.93%	128.79%
1 level of AMR	21,878	65,634	29,894	45.55%	136.64%
2 levels of AMR	54,234	162,702	81,502	50.09%	150.28%

Table 9

Cost-efficiency of the proposed mesh adaptivity scheme for the Bocca beam.

	Number of elements with initial uniform refinement	Max. number of elements with AMR	%
No mesh refinement	6,464	–	–
1 level of refinement	25,856	10,556	40.83%
2 levels of refinement	103,424	26,402	25.53%

Table 10

CPU-times of the computations with standard FEs, mixed FEs and the proposed AFR scheme for the Bocca beam.

	CPU-times (s)		
	Standard FEs	AFR	Mixed FEs
Without AMR	757.09	1,193.07	5,248.93
1 level of AMR	4,091.41	4,390.56	18,556.28
2 levels of AMR	12,490.85	25,394.72	76,053.59

opening displacement (CMOD) control of the upper notch.

In order to assess the proposed AMR and AFR capabilities, a coarse initial mesh is considered for the present case study, and different levels

of adaptive mesh refinement are subsequently introduced. The initial mesh, shown in Fig. 25, has a FE size of 5 mm, which results in a total of 6,464 elements. This section intends to analyze the capabilities of the model when more than one crack develops.

In this case study, mesh refinement takes place when the equivalent effective stress $\bar{\sigma}_{eq} \geq 0.5f_t$ and coarsening occurs when $\bar{\sigma}_{eq} \leq 0.25f_t$. The mixed FE formulation is activated when $\bar{\sigma}_{eq} \geq 0.75f_t$ and deactivated when $\bar{\sigma}_{eq} \leq 0.25f_t$.

Fig. 26 shows the computed displacement contours of the Bocca beam, where results (i) without AMR, (ii) with 1 level of AMR and (iii) with 2 levels of AMR are contrasted. The position of the cracks can be located due to the high gradient of displacements that they generated. It can be observed how the beam breaks into 3 different pieces, and that the central piece is exhibiting a pure rotation motion, due to the polar symmetry of the applied boundary conditions. Note how the displacement gradients are much more localized when a lower FE size is adopted, due to the adaptive mesh refinement capabilities of the model.

In Fig. 27 the computed damage contours of the Bocca beam are shown, comparing results (i) without AMR, (ii) with 1 level of AMR and (iii) with 2 levels of AMR. It can be seen how two cracks with polar symmetry are developing in the beam, as is expected from the applied boundary conditions. Fig. 28 examines the detail of the computed crack trajectories. It can be observed how the three simulations are producing practically overlapping results. Also, note that the computed fractures are inside the range of experimental observations.

Fig. 29 shows the results in terms of force F vs displacement curves at the upper right support. The three computations considered produce practically matching results, with a small overestimation of the bearing capacity in the case calculated with the coarse mesh. It can be observed how the simulations with 1 and 2 levels of AMR provide almost the same peak load. Fig. 29 also details how the computed force–displacement curves agree with the experiment. This is remarkable, given that only one curve is provided for the original experiment in reference [57].

In Fig. 30 the evolution of the areas where standard and mixed FE formulations are active for different time-steps is shown. It can be seen how the proposed criterion defined in function of the equivalent effective stress is able of successfully delimiting the areas where mixed FEs should be activated in the considered case study where two different cracks are evolving. Fig. 31 presents the meshes that result at the end of

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

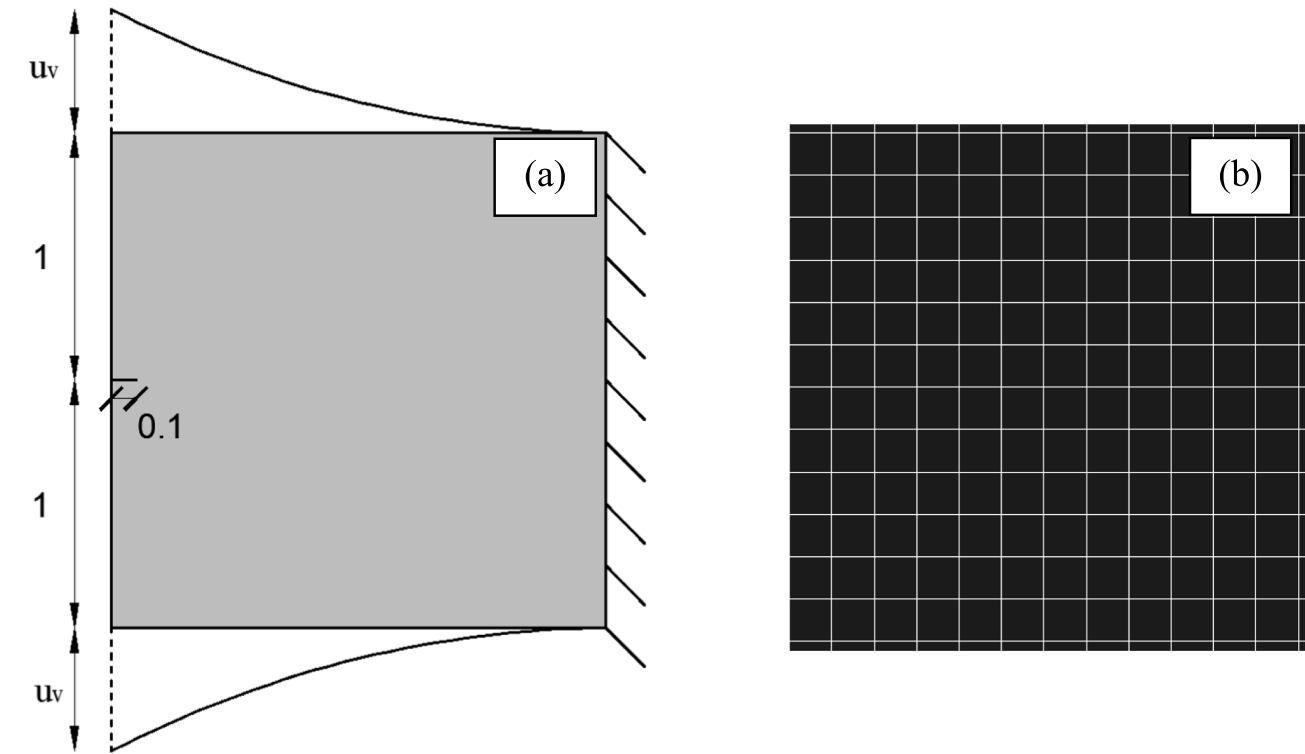


Fig. 32. Crack branching numerical test: (a) Geometry (mm) and (b) detail of the mesh in the area of the notch.

Table 11
Material properties of the crack branching numerical test.

Young's Modulus	20.0·10 ⁹ Pa
Poisson's Ratio	0.3
Tensile Strength	44.0·10 ⁶ Pa
Tensile Fracture Energy	89 J/m ²

the simulation when the adaptive mesh refinement capabilities are introduced. It can be seen how the proposed model is also able to properly define the regions that required mesh refinement and later coarsening.

In Table 8 the cost-efficiency of the proposed methodology is assessed. As in the previous cases, for each of the situations considered (without AMR, with 1 level of AMR and with 2 levels of AMR) the maximum number of degrees of freedom is indicated for analyzes performed with standard FE, mixed FE and AFR. The relation between the

maximum number of degrees of freedom in AFR and the other two methods (standard only and mixed only) is also given. In the present analysis, where an initial coarse mesh is considered, it can be observed how the reduction in computational cost provided by AFR is not as important as in the previous sections where an initially fine mesh is adopted. Notwithstanding, the use of the AFR approach remains highly advantageous also in this instance.

Table 9 shows the cost-effectiveness of AMR in the Bocca beam analysis. For the three cases considered (without refinement, with 1 level of refinement and with 2 levels of refinement) the strategies of uniform refinement (first column) and adaptive refinement (second column) are compared in terms of the number of elements that ensue in the computation. In the third column the ratio between the two values is indicated. It can be observed how the AMR capabilities of the model provide a noteworthy reduction in computational cost.

To further show the increased efficiency ensuing from the adoption of the proposed AFR scheme, Table 10 compares the resulting CPU-times

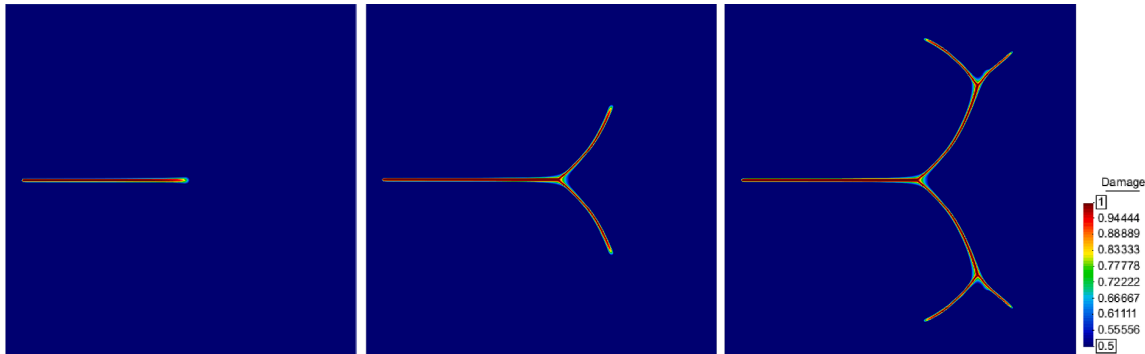


Fig. 33. Computed damage contours at different stages of the simulation of the crack branching numerical test.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

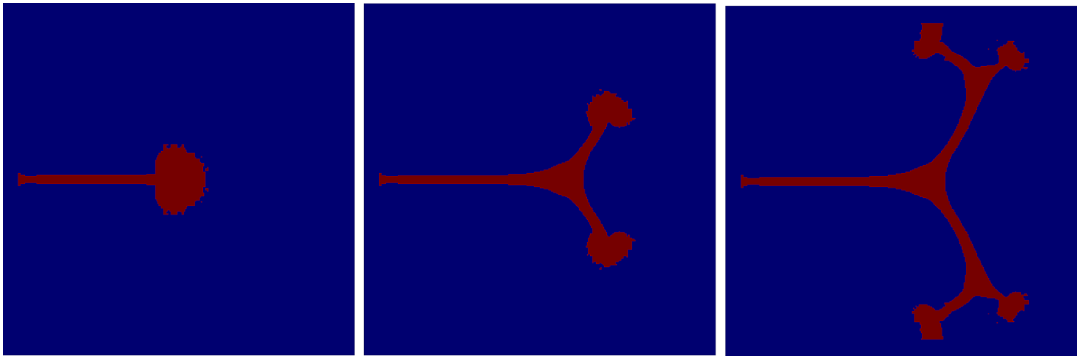


Fig. 34. Crack branching numerical test: Evolution of the area where standard (in blue) and mixed (in red) formulations are adopted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

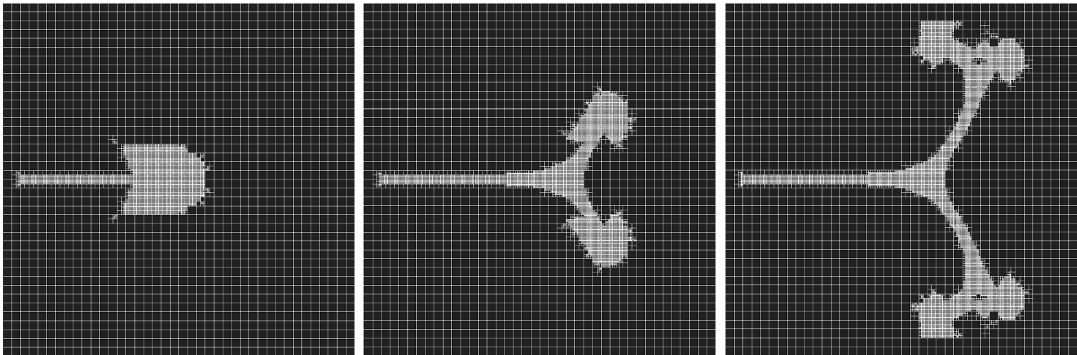


Fig. 35. Crack branching numerical test: Evolution of the refined area.

Table 12
Cost-efficiency of the proposed formulation adaptivity scheme for the crack branching numerical test.

Max. Ndots mixed ϵ/u FEs	Max. Ndots AFR	% of Ndots with respect to mixed only
87,132	66,332	76.13%

Table 13
Cost-efficiency of the proposed mesh adaptivity scheme for the crack branching numerical test.

Number of nodes with initial uniform refinement	Max. number of nodes with adaptive mesh refinement	%
103,057	14,522	14.09

with corresponding simulations using standard and mixed FEs only. Again, each of the considered cases (without AMR, with 1 level of AMR and with 2 levels of AMR) are considered. Computations were carried out in a single thread on an Intel Core i7-6700 CPU. It can be observed that the CPU-times ensuing from the application of the AFR strategy are much lower than the ones corresponding to the use of mixed FEs. The order of magnitude of the computational cost of AFR is comparable to the one resulting from the use of standard FEs. This shows that the AFR approach highly reduces the cost of the simulations performed with mixed FEs, while preserving the quality of the results.

4.4. Crack branching numerical test

In this section, the numerical simulation of a crack branching test is

performed. In this benchmark, initially proposed in reference [64], a square plate with an initial notch at the midheight is considered in order to examine the ability of the numerical approach in reproducing bifurcations.

The geometry and boundary conditions of the problem are indicated in Fig. 32a. The plate has dimensions of 2 mm × 2 mm. The thickness of the notch is taken as zero. The right edge of the plate is fixed and on the top and bottom boundaries are imposed vertical displacements following the function $f(x) = u_v(x - 1)^2/4$, u_v being the displacement at the top-left and bottom-left corners and the origin $x = 0$ being placed at the center of the specimen so that $-1 \leq x \leq 1$. These imposed boundary conditions enforce the branching of the propagating crack in the plate. The material properties of the test are included in Table 11, used together with the Rankine damage surface.

The numerical simulation is performed with both AFR and AMR capabilities. In this case, 3 levels of mesh refinement are introduced. AMR is triggered when the equivalent effective stress $\bar{\sigma}_{eq}$ reaches 75% of the tensile strength f_t . Coarsening is enabled when $\bar{\sigma}_{eq}$ drops below 50% of f_t . The same thresholds are established for AFR.

The detail of the initial mesh adopted to perform the simulation in the area of the notch is shown in Fig. 32b. It is a structured mesh of $40 \times 40 = 1600$ quadrilateral elements. Note that this original mesh is not fine enough to produce an acceptable result of the numerical simulation if AMR is not introduced. It has been chosen to illustrate the capabilities of the proposed numerical scheme when a coarse mesh is initially considered.

In Fig. 33 the computed damage contours at different stages of the simulation can be observed. It can be seen how first the crack progresses horizontally and then branches due to the boundary conditions imposed in the body. Two symmetric cracks evolve in the plate until branching occurs again. Then, four symmetric cracks propagate through the plate.

G.B. Barbat et al.

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

Fig. 34 shows the evolution of the area where the mixed FE formulation is activated during the simulation. It can be seen how mixed FEs are introduced in the simulation only in the areas where the cracks propagate. Also, in Fig. 35 the regions where the mesh is refined can be observed for the same time steps. This shows that the proposed criterion based on the equivalent effective stress is satisfactorily defining the areas to be refined in crack branching situations as well.

Table 12 assesses the savings in computational cost due to the proposed AFR scheme. The first column indicates the maximum number of degrees of freedom (labelled Max. Ndots) ensuing from computations with mixed FE only while the second column shows the corresponding value for the AFR strategy. The last column presents the relation between the two previous values. In the present case where an initial coarse mesh is adopted, the decrease in the number of degrees of freedom due to the AFR scheme is not as important as in the previous case studies where an initially fine mesh was employed. Still, the FE formulation adaptivity allows to introduce a significant reduction in computational cost.

Table 13 examines the cost-efficiency of the adopted AMR strategy. In the first column the number of nodes resulting from the use of an initial uniform refinement approach is provided while in the second column the number of nodes introduced by the ARM method is indicated. The third column shows the relation between the two previous values. It can be inferred that in this case where an initial coarse mesh is adopted to calculate the problem, AMR produces a large saving in the computational cost with respect to the use of initial mesh refinement.

In the present case study, the adaptive mesh refinement strategy produces a large part of the total cost savings of the method. No sign of spurious mesh bias dependence can be observed. The combination of the AMR and AFR capabilities of the proposed model allows to obtain mesh objective results at a reasonable computational cost.

5. Conclusions

In this work, the use of an Adaptive Formulation Refinement (AFR) strategy is combined with the application of an Adaptive Mesh Refinement (AMR) approach for the cost-efficient computation of localized structural failure with enhanced accuracy and mesh objectivity.

The proposed AFR strategy allows to adaptively switch between the standard displacement-based FE formulation and the more accurate mixed strain/displacement FE. An isotropic damage model is used to model the stiffness degradation leading to cracking and structural failure. A simple and effective criterion defined in terms of the ratio between the equivalent effective stresses that develop in the structure and the material strength is introduced to selectively activate the formulations. The proposed methodology has shown to significantly reduce the computational cost of the simulations, rendering them of the order of standard FE calculations, while benefitting from the enhanced accuracy and mesh bias objectivity of the mixed FE.

Using a similar refinement criterion based on the same ratio between the equivalent effective stresses and the material strength, the adopted AMR strategy further reduces the required computational cost of structural analyses, by allowing to start the simulation with a coarse mesh and adaptively refining only the areas where crack propagation occurs, which represent a small part of the whole domain.

The proposed model also allows for the later adaptive mesh coarsening and *formulation derefinement*, switching back from mixed to standard FEs, according to the corresponding criteria. The combination of the AFR and AMR methodologies is fairly straightforward.

The paper includes the assessment of the proposed methodology through the simulation of several numerical benchmarks and experimental tests available in the literature. The successful computation of situations involving structural size effect, multiple developing cracks and crack branching is presented. The simulations reproduce the bearing capacity, force-displacement curves, crack trajectories and collapse mechanism observed in the experiments with enhanced accuracy.

Complementarily, the comparison of the proposed model with the standard displacement-based formulation is also performed. This shows, on the one hand, the capacity of the adopted methodology combining AFR and AMR to produce converging mesh objective results, without any trace of spurious mesh dependence. On the other hand, the standard formulation alone remains spuriously mesh dependent and unreliable, even if the mesh is refined.

Calculations show the capacity of the approach integrating AFR and AMR to obtain reliable results with high accuracy and with a computational cost of the order of the standard displacement-based FE. This demonstrates the effectiveness of the approach in engineering practice.

CRedit authorship contribution statement

G.B. Barbat: Software, Formal analysis, Writing – original draft. **M. Cervera:** Conceptualization, Supervision, Writing – review & editing. **H. Venghaus:** Methodology, Software. **M. Chiumenti:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The finite element code employed is not open source.

Acknowledgements

Financial support from the European Union's horizon 2020 research and innovation programme (H2020-DT-2019-1 No. 872570) under the KYKLOS 4.0 Project (An Advanced Circular and Agile Manufacturing Ecosystem based on rapid reconfigurable manufacturing process and individualized consumer preferences) and by the Spanish Ministry of Science, Innovation and Universities via the Primus project (Computational Framework for Additive Manufacturing of Titanium Alloy) project (Printing pattern based and MultiScale enhanced performance analysis of advanced Additive Manufacturing components, ref. num. PID2020-115575RB-I00) is gratefully acknowledged. The authors also acknowledge financial support from the Spanish Ministry of Economy and Competitiveness, through the "Severo Ochoa Programme for Centres of Excellence in R&D" (CEX2018-000797-S).

References

- [1] J. Rots, P. Nauta, G. Kusters, J. Blauwendraad, Smeared crack approach and fracture localization in concrete, *Heron* 30 (1985) 1–48. <http://resolver.tudelft.nl/uuid:746aba89-3566-4ade-9931-d6bbae0c05ea>.
- [2] J. Rots, Computational modeling of concrete fracture, Ph.D. Thesis, Delft University of Technology, 1988. <http://resolver.tudelft.nl/uuid:06985d0d-1230-4a08-924a-2553a171f08f>.
- [3] J. Oliver, M. Cervera, S. Oller and J. Lubliner, Isotropic damage models and smeared crack analysis of concrete, II Int. conference on computer aided analysis and design of concrete, pp. 945–957, 1990.
- [4] L. Sluys, Wave propagation localisation and dispersion in softening solids, Ph.D. Thesis, Delft University of Technology, 1992. <http://resolver.tudelft.nl/uuid:0f9b3de9-e0ec-4d9b-a42f-fcf704d5d40e>.
- [5] T. Rabczuk, Computational methods for fracture in brittle and quasi-brittle solids: State-of-the-art review and future perspectives, *ISRN Appl. Math.* 2013 (2013), 849231, <https://doi.org/10.1155/2013/849231>.
- [6] M. Cervera, G. Barbat, M. Chiumenti, J.-Y. Wu, A Comparative Review of XFEM, Mixed FEM and Phase-Field Models for Quasi-brittle Cracking, *Arch. Comput. Methods Eng.* 29 (2022) 1009–1083.
- [7] M. Cervera, M. Chiumenti, R. Codina, Mixed stabilized finite element methods in nonlinear solid mechanics. Part I: Formulation, *Comput. Methods Appl. Mech. Eng.* 199 (37–40) (2010) 2559–2570.
- [8] M. Cervera, M. Chiumenti, R. Codina, Mixed stabilized finite element methods in nonlinear solid mechanics. Part II: Strain localization, *Comput. Methods Appl. Mech. Eng.* 199 (37–40) (2010) 2571–2589.

G.B. Barbat *et al.*

Theoretical and Applied Fracture Mechanics 122 (2022) 103646

- [9] M. Cervera, M. Chiumenti, Smeared crack approach: back to the original track, *Int. J. Numer. Anal. Meth. Geomech.* 30 (12) (2006) 1173–1199.
- [10] M. Cervera, M. Chiumenti, R. Codina, Mesh objective modeling of cracks using continuous linear strain and displacement interpolations, *Int. J. Numer. Meth. Eng.* 87 (10) (2011) 962–987.
- [11] M. Cervera, G. Barbat, M. Chiumenti, Finite element modelling of quasi-brittle cracks in 2D and 3D with enhanced strain accuracy, *Comput. Mech.* 60 (5) (2017) 767–796.
- [12] G. Barbat, M. Cervera, M. Chiumenti, Appraisal of planar, bending and twisting cracks in 3D with isotropic and orthotropic damage models, *Int. J. Fract.* 210 (1–2) (2018) 45–79.
- [13] G. Vlachakis, M. Cervera, G. Barbat, S. Saloustros, Out-of-plane seismic response and failure mechanism of masonry structures using finite elements with enhanced strain accuracy, *Engineering Failure Analysis* 97 (2019) 534–555.
- [14] M. Cervera, G. Barbat, M. Chiumenti, Architecture of a multi-crack model with full closing, reopening and sliding capabilities, *Comput. Mech.* 65 (6) (2020) 1593–1620.
- [15] G. Barbat, M. Cervera, M. Chiumenti, E. Espinoza, Structural size effect: experimental, theoretical and accurate computational assessment, *Eng. Struct.* 213 (2020), 110555.
- [16] C. Agelet de Saracibar, M. Chiumenti, Q. Valverde, M. Cervera, On the orthogonal subgrid scale pressure stabilization of finite deformation J2 plasticity, *Comput. Methods Appl. Mech. Eng.* 195 (9–12) (2006) 1224–1251.
- [17] M. Chiumenti, M. Cervera, R. Codina, A mixed three-field FE formulation for stress accurate analysis including the incompressible limit, *Comput. Methods Appl. Mech. Eng.* 283 (2015) 1095–1116.
- [18] M. Chiumenti, M. Cervera, C. Moreira, G. Barbat, Stress, strain and dissipation accurate 3-field formulation for inelastic isochoric deformation, *Finite Elem. Anal. Des.* 192 (2021), 103534.
- [19] C. Moreira, G. Barbat, M. Cervera, M. Chiumenti, Accurate thermal-induced structural failure analysis under incompressible conditions, *Eng. Struct.* 261 (2022), 114213.
- [20] Inocencio Castañar, Joan Baiges, Ramon Codina, A stabilized mixed finite element approximation for incompressible finite strain solid dynamics using a total Lagrangian formulation, *Computer Methods in Applied Mechanics and Engineering*, Volume 368, 2020, 113164, ISSN 0045-7825, <https://doi.org/10.1016/j.cma.2020.113164>.
- [21] Inocencio Castañar, Joan Baiges, Ramon Codina, Henning Venghaus, Topological derivative-based topology optimization of incompressible structures using mixed formulations, *Computer Methods in Applied Mechanics and Engineering*, Volume 390, 2022, 114438, ISSN 0045-7825, <https://doi.org/10.1016/j.cma.2021.114438>.
- [22] H. Venghaus, M. Chiumenti, J. Baiges, D. Juhre and I. Castañar, An accurate approach to simulate the friction stir welding process using adaptive formulation refinement, submitted to CAME, 2022.
- [23] M. Borden, C. Verhoosel, M. Scott, T. Hughes, C. Landis, A phase-field description of dynamic brittle fracture, *Comput. Methods Appl. Mech. Eng.* 217–220 (2012) 77–95.
- [24] T. Heister, M. Wheeler, T. Wick, A primal-dual active set method and predictor-corrector mesh adaptivity for computing fracture propagation using a phase-field approach, *Comput. Methods Appl. Mech. Eng.* 290 (2015) 466–495.
- [25] S. Zhou, X. Zhuang, Adaptive phase field simulation of quasi-static crack propagation in rocks, *Underground Space* 3 (3) (2018) 190–205.
- [26] N. Ferro, S. Micheletti, S. Perotto, Anisotropic mesh adaptation for crack propagation induced by a thermal shock in 2D, *Comput. Methods Appl. Mech. Eng.* 331 (2018) 138–158.
- [27] C. Hirshikesh, K. Jansari, R.A. Kannan, S. Natarajan, Adaptive phase field method for quasi-static brittle fracture using a recovery based error indicator and quadtree decomposition, *Eng. Fract. Mech.* 220 (2019), 106599, <https://doi.org/10.1016/j.engfracmech.2019.106599>.
- [28] J. Baiges, C. Bayona, Refficientlib: An efficient load-rebalanced adaptive mesh refinement algorithm for high-performance computational physics meshes, *SIAM J. Scientific Comput.* 39 (2) (2017) 65–95, <https://doi.org/10.1137/15M105330X>.
- [29] J. Baiges, R. Codina, A. Pont, E. Castillo, An adaptive fixed-mesh ALE method for free surface flows, *Comput. Methods Appl. Mech. Eng.* 313 (2017) 159–188.
- [30] J. Baiges, M. Chiumenti, C. Moreira, M. Cervera, R. Codina, An adaptive finite element strategy for the numerical simulation of additive manufacturing processes, *Addit. Manuf.* 37 (2021), 101650.
- [31] C. Moreira, M. Caicedo, M. Cervera, M. Chiumenti, J. Baiges, A multi-criteria h-adaptive finite element framework for industrial part-scale thermal analysis in additive manufacturing processes, *Eng. Comput.* (2022), <https://doi.org/10.1007/s00366-022-01655-0>.
- [32] Y. Rashid, Ultimate strength analysis of prestressed concrete pressure vessels, *Nucl. Eng. Des.* 7 (4) (1968) 334–344.
- [33] Z. Bazant, B. Oh, Crack band theory for fracture of concrete, *Matériaux et Constructions* 16 (3) (1983) 155–177, <https://doi.org/10.1007/BF02486267>.
- [34] J.-Y. Wu, J.-F. Qiu, V. Nguyen, T. Mandal, L.-J. Zhuang, Computational modeling of localized failure in solids: XFEM vs PF-CZM, *Comput. Methods Appl. Mech. Eng.* 345 (2019) 618–643.
- [35] M. Cervera, M. Chiumenti, L. Benedetti, R. Codina, Mixed stabilized finite element methods in nonlinear solid mechanics. Part III: Compressible and incompressible plasticity, *Comput. Methods Appl. Mech. Eng.* 285 (2015) 752–775.
- [36] L. Benedetti, M. Cervera, M. Chiumenti, 3D modelling of twisting cracks under bending and torsion skew notched beams, *Eng. Fract. Mech.* 176 (2017) 235–256.
- [37] S. Saloustros, M. Cervera, S. Kim, M. Chiumenti, Accurate and locking-free analysis of beams, plates and shells using solid elements, *Comput. Mech.* 67 (2021) 883–914.
- [38] F. Brezzi, K. Bathe, A Discourse on the Stability Conditions for Mixed Finite Element Formulations, *Comput. Methods Appl. Mech. Eng.* 82 (1990) 27–59, [https://doi.org/10.1016/0045-7825\(90\)90157-H](https://doi.org/10.1016/0045-7825(90)90157-H).
- [39] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers, *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique* 8 (R2) (1974) 129–151. <https://eudml.org/doc/193255>.
- [40] F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer, New York, 1991, <https://doi.org/10.1007/978-1-4612-3172-1>.
- [41] I. Babuska, Error-bounds for finite element method, *Numerisch Mathematik* 16 (1971) 322–333, <https://doi.org/10.1007/BF02165003>.
- [42] D. Boffi, F. Brezzi, M. Fortin, *Mixed Finite Element Methods and Applications*, Springer, 2013, <https://doi.org/10.1007/978-3-642-36519-5>.
- [43] R. Codina, Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods, *Comput. Methods Appl. Mech. Eng.* 190 (2000) 1579–1599.
- [44] T. Hughes, Multiscale phenomena: Green’s function, Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations, *Comput. Methods Appl. Mech. Eng.* 127 (1–4) (1995) 387–401.
- [45] R. Codina, Analysis of a stabilized finite element approximation of the Oseen equations using orthogonal subscales, *Appl. Numer. Math.* 58 (2008) 264–283.
- [46] R. Codina, Finite element approximation of the three field formulation of the Stokes problem using arbitrary interpolations, *Soc. Indust. Appl. Math. J. Numer. Anal.* 47 (2009) 699–718.
- [47] S. Badia, R. Codina, Unified stabilized finite element formulations for the Stokes and the Darcy problems, *SIAM J. Numer. Anal.* 47 (2009) 1971–2000, <https://doi.org/10.1137/08072632X>.
- [48] L. Benedetti, M. Cervera, M. Chiumenti, High-fidelity prediction of crack formation in 2D and 3D pullout tests, *Comput. Struct.* 172 (2016) 93–109.
- [49] L. Benedetti, M. Cervera, M. Chiumenti, Stress-accurate mixed FEM for soil failure under shallow foundations involving strain localization in plasticity, *Comput. Geotech.* 64 (2015) 32–47.
- [50] M. Cervera, J.-Y. Wu, M. Chiumenti, S. Kim, Strain localization analysis of Hill’s orthotropic elastoplasticity: Analytical results and numerical verification, *Comput. Mech.* 65 (2019) 533–554, <https://doi.org/10.1007/s00466-019-01782-4>.
- [51] L. Benedetti, Mixed finite element formulations for strain localization and failure in plasticity, Ph.D. Thesis, Technical University of Catalonia, 2017. <http://hdl.handle.net/2117/108236>.
- [52] A. Coll, R. Ribó, M. Pasenau, E. Escolano, J. S. Perez, A. Melendo, A. Monros and J. Garate, GiD: the personal pre and post-processor User Manual, CIMNE, Technical University of Catalonia, <http://gid.cimne.upc.edu>, 2018.
- [53] V. Garcia-Alvarez, R. Gettu, I. Carol, Analysis of mixed-mode fracture in concrete using interface elements and a cohesive crack model, *Sadhana* 37 (1) (2012) 187–205.
- [54] D.-C. Feng, J.-Y. Wu, Phase-field regularized cohesive zone model (CZM) and size effect of concrete, *Eng. Fract. Mech.* 197 (2018) 66–79.
- [55] G. Barbat, Mixed finite elements with independent strain interpolation for isotropic and orthotropic damage, Technical University of Catalonia, 2021. Ph.D. Thesis, <http://hdl.handle.net/2117/346007>.
- [56] M. Arrea and A. Ingraffea, “Mixed-mode crack propagation in mortar and concrete,” Tech. Rep. Report No. 81-13, Department of Structural Engineering, Cornell University, New York, 1982.
- [57] P. Bocca, A. Carpinteri, S. Valente, Mixed mode fracture of concrete, *Int. J. Solids Struct.* 27 (9) (1991) 1139–1153.
- [58] T. Rabczuk, T. Belytschko, Cracking particles: a simplified meshfree method for arbitrary evolving cracks, *Int. J. Numer. Meth. Eng.* 61 (2004) 2316–2343, <https://doi.org/10.1002/nme.1151>.
- [59] T. Rabczuk, G. Zi, A meshfree method based on the local partition of unity for cohesive cracks, *Comput. Mech.* 39 (2006) 743–760, <https://doi.org/10.1007/s00466-006-0067-4>.
- [60] A. Saleh, M. Aliabadi, Crack growth analysis in concrete using boundary element method, *Eng. Fract. Mech.* 51 (4) (1995) 533–545.
- [61] Y. Zhang, R. Lackner, M. Zeiml, H. Mang, Strong discontinuity embedded approach with standard SOS formulation: element formulation, energy-based crack-tracking strategy, and validations, *Comput. Methods Appl. Mech. Eng.* 287 (2015) 335–366.
- [62] P. Areias, T. Rabczuk, D. Dias-da-Costa, Element-wise fracture algorithm based on rotation of edges, *Eng. Fract. Mech.* 110 (2013) 113–137.
- [63] P. Areias, T. Rabczuk, J. Cesar de Sa, A novel two-stage discrete crack method based on the screened Poisson equation and local mesh refinement, *Comput. Mech.* 58 (2016) 1003–1018.
- [64] A. Muixi, O. Marco, A. Rodriguez-Ferran, S. Fernandez-Mendez, A combined XFEM phase-field computational model for crack growth without remeshing, *Comput. Mech.* 67 (2021) 231–249.

Publication 3

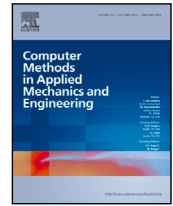
C



Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma



Embedded technology for enhanced modeling of Friction Stir Welding processes

H. Venghaus^{a,b,*}, M. Chiumenti^{a,b}, J. Baiges^{a,b}, D. Juhre^c, N. Dialami^{a,b}

^a Polytechnic University of Catalonia (UPC), Barcelona Tech, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain

^b International Center for Numerical Methods in Engineering (CIMNE), Polytechnic University of Catalonia (UPC), Campus Norte UPC, 08034 Barcelona, Spain

^c Institute of Mechanics (IFME), Otto von Guericke University Magdeburg (OVGU), Universitätsplatz 2, 39106 Magdeburg, Germany

ARTICLE INFO

Keywords:

Friction Stir Welding
Finite element method
Embedded domain method
Cut finite elements
Experimental validation

ABSTRACT

Friction Stir Welding (FSW) is a solid-state joining process that has several benefits over conventional welding techniques. A major challenge is its high sensitivity to process parameters such as advancing and rotational speed. Simulations are a key tool for understanding the material flow and temperature evolution and help to find the best processing parameters for a given FSW task. This work proposes using the Embedded Domain Method (EDM) for the simulation of FSW to overcome the limitations of (purely) Lagrangian, Eulerian, or Arbitrary Lagrangian Eulerian (ALE) approaches.

In the proposed approach, a CAD geometry of the rotating pin tool is embedded at runtime into the computational domain, which represents the to-be-welded workpieces. Resulting interface elements are integrated using a boundary conforming subintegration approach, and various stabilization techniques are employed to stabilize the thermo-mechanical problem and potential cut element instabilities.

The predicted reaction forces and temperature evolution curves are in excellent agreement with experimental reference data and the sensitivity analysis of process parameters shows that the model can be used to make targeted adjustments of said parameters to gain greater control over the process.

1. Introduction

Friction Stir Welding (FSW) is a solid-state joining process in which two facing workpieces (Fig. 1a) are welded together using a rotating, non-consumable tool (Fig. 1b) that is plunged into the material and then moved along their interface. In FSW, the joining process is achieved because of the heat generated by friction between the pin-tool shoulder and the workpieces, as well as by the plastic deformation of the stirred material. Typically, the material is not melted.

This is a great advantage over conventional welding techniques since any cracking due to phase change is avoided, and welding can be conducted in arbitrary orientations. Despite the advantages and significant attention FSW has received within industrial sectors (especially aerospace and naval), its complexity lies in its high sensitivity to various process parameters. Controlling factors such as rotation speed, advancing speed, pin shape, and tool tilting are key to obtaining high-quality, defect-free welds. The numerical simulation of FSW processes can help understand these parameters' influence on the weld quality (see [2]). Simulations

* Corresponding author.

E-mail address: henning.venghaus@upc.edu (H. Venghaus).

<https://doi.org/10.1016/j.cma.2024.117539>

Received 5 August 2024; Received in revised form 20 October 2024; Accepted 6 November 2024

Available online 30 November 2024

0045-7825/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

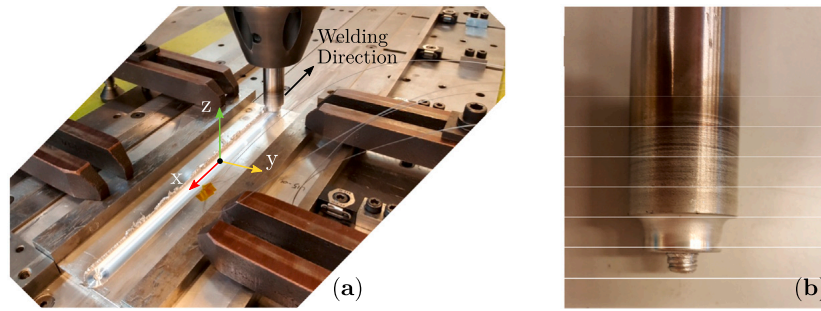


Fig. 1. (a) Experimental welding setup, and (b) pin tool, [1].

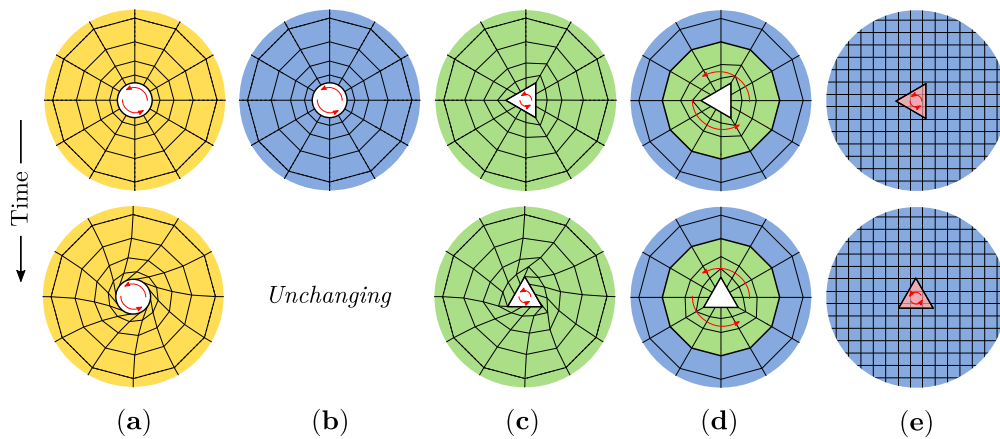


Fig. 2. Different approaches to simulating FSW: (a) entirely Lagrangian frame of reference, (b) entirely Eulerian frame of reference, (c) entirely ALE frame of reference, (d) AKF, partially Eulerian, partially ALE, and (e) EDM approach with Eulerian frame of reference and embedded domain. Colorcoding: ■ Lagrangian, ■ Eulerian, ■ ALE, ■ Embedded Domain.

of FSW processes are mainly focused on the prediction of the temperature distribution [3,4], the sensitivity to the process parameters [5], the analysis of the strain rate field [6] and the material flow [7,8] as well as the contact and friction [9,10].

Due to the high temperatures in FSW, the workpiece material softens significantly in close vicinity to the tool. Considering this, researchers have used different kinematic frameworks to model FSW processes. In a Lagrangian framework, where the computational mesh is fixed to material particles (Fig. 2a), the workpiece is modeled as a visco-elasto-plastic solid, and the primary unknowns are displacements. This approach allows for a very accurate description of the material's movement and, most importantly, a residual stress analysis. However, significant drawbacks are excessive mesh distortions, which lead to frequent and costly re-meshing operations, including transferring unknowns and internal variables between meshes.

In contrast, there are Eulerian frameworks, where the mesh is fixed in space, the workpiece is characterized as a (very viscous) fluid, and the primary unknowns are velocities. In an Eulerian framework, simpler material and kinematic models can be used (visco-plastic stokes flow vs. finite strain elasto-visco-plastic solid mechanics), and faster solution times [11] can be achieved.

In the particular case of rotationally symmetric tools, the entire simulation can be carried out using an unchanging mesh (Fig. 2b). In this case, the simulation converges to a steady-state solution. However, non-rotationally symmetric tools, also require re-meshing; thus, an Arbitrary-Lagrangian-Eulerian (ALE) formulation can be an alternative (Fig. 2c).

A proposed solution to this problem is the Apropos Kinematic Framework (AKF, Fig. 2d) by Dialami et al. [12], which divides the domain into a circular ALE subdomain that rotates in unison with the tool and the far-field that is described by an Eulerian framework. Special attention must be given to the interface between the two subdomains. Consequently, the AKF allows for the analysis of non-rotationally symmetric tools while avoiding any re-meshing and capitalizes on the benefits of the Eulerian framework. Unfortunately, the AKF requires the pin rotation axis $\mathbf{k}$ to be parallel to the workpiece surface normal $\mathbf{n}$. If this condition is not fulfilled, the approach cannot be used, as demonstrated in Fig. 3. This complicates the inclusion of tilt angles.

To avoid this problem, the authors present a new approach to simulate Eulerian FSW problems using an Embedded Domain Method (EDM, Fig. 2e). EDMs are a group of closely related methods characterized by an unfitted mesh and (material) discontinuities within elements. The most prominent examples are the *Extended Finite Element Method* (XFEM, see [13,14]), which enriches the solution space with discontinuous functions, the *Finite Cell Method* (FCM, see [15,16]), which is characterized by large, unfitted elements called cells and high order shape functions, the *Immersed Boundary Method* (IBM, see [17]) which focuses on defining boundary conditions independently from mesh and element boundaries and the *Fictitious Domain Method* (FDM, see [18,19]), which

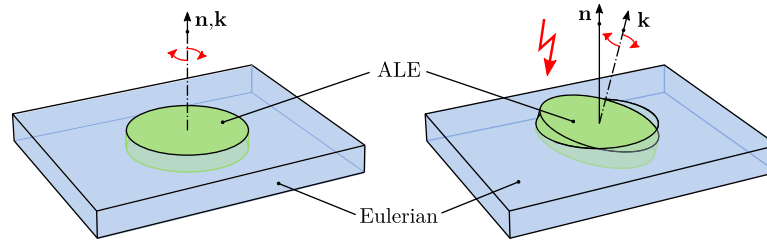


Fig. 3. Axis misalignment encountered in Apropos Kinematic Framework for non-zero tilt angles.

assumes either the embedded domain or its complementary part fictitious and thus not contributing to the solution. However, most mentioned methods lack clear differentiation, and may overlap in certain aspects.

Janoske et al. [20] have used the IBM to simulate FSW by immersing an STL geometry of a pin-tool into a computational fluid dynamics (CFD) domain. This approach allows one to work with unfitted meshes and to impose the problem's boundary conditions without respecting the element's boundaries. IBM, however, does not differentiate between the *inside* and *outside* of the immersed tool, influencing the overall solution as the workpiece and tool subdomains remain coupled.

The present work follows a similar approach by performing the computation on a regular mesh that is also not body-fitted. The FSW tool is *embedded* into the computational domain (workpiece geometry). However, unlike in the IBM approach, elements are classified as *physical* (not intersected by the tool), *inactive* (entirely inside the tool subdomain), or *cut* by the tool boundary. When integrating over the domain, the physical elements are treated as usual, the inactive ones do not contribute to the solution, and the cut elements receive special treatment to account for their partial contribution to the physical subdomain. This approach effectively decouples the workpiece and tool subdomains, eliminating all interactions. Special numerical stabilization techniques are applied to deal with the numerical instabilities and negatively impacted system conditioning caused by small cut elements. Finally, the friction condition is applied to the tool-workpiece interface, which no longer coincides with the element boundaries. This approach is best described as CUT-FEM with low-order elements, see [21].

This approach allows for using complex, non-rotationally symmetric tools while simultaneously allowing for non-zero tilt angles. Furthermore, the workflow from tool design (using CAD software) to Finite Element simulation is fully automated and does not require any manual mesh generation and verification. Finally, an octree refinement strategy is employed, hierarchically refining the mesh close to the tool-workpiece interface. This refinement simultaneously serves two purposes (accurate geometry capture and solution enhancement in the boundary layer) without needing a refinement criterium based on elaborate error estimation.

The outline of the paper is as follows: Section 2 summarizes the main equations of the thermo-mechanically coupled problem and the material model and gives a detailed description of the problem's boundary conditions, Section 3 covers the chosen EDM approach and outlines the *CAD-to-Simulation* workflow and finally Sections 4 and 5 show selected numerical results and draw conclusions.

2. Thermo-mechanical analysis

This section presents the fundamental equations of the thermo-mechanical problem in their strong and weak forms, as well as various stabilization terms. Furthermore the boundary conditions and the material model are explained.

2.1. Governing equations

Friction Stir Welding is a two-way, thermo-mechanically coupled process, as summarized in Fig. 4. The mechanical problem is modeled as an incompressible Stokes flow with a non-Newtonian (shear-thinning) and temperature-dependent constitutive law. The thermal problem describes the conduction and convection of heat generated by plastic dissipation and friction from the mechanical problem. The thermal problem is convection-dominated, and the convective velocity is one of the unknowns of the mechanical problem. The coupled problem is solved in a staggered manner. A detailed discussion of the governing equations of FSW and the derivation of the weak form can be found in [2] and [22].

Mechanical. The mechanical problem is described by the balance of momentum, Eq. (1), and the incompressibility constraint, Eq. (2).

$$\nabla \cdot \mathbf{s} + \nabla p + \mathbf{b} = \mathbf{0} \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (2)$$

The problem is expressed in terms of two primary unknowns velocity $\mathbf{u}$ and pressure p , as well as the deviatoric stress tensor $\mathbf{s}$ (see Section 2.3) and the body forces $\mathbf{b}$. The problem is presented in *mixed* $\mathbf{u}/p$ form to deal with the isochoric nature of the material flow.

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

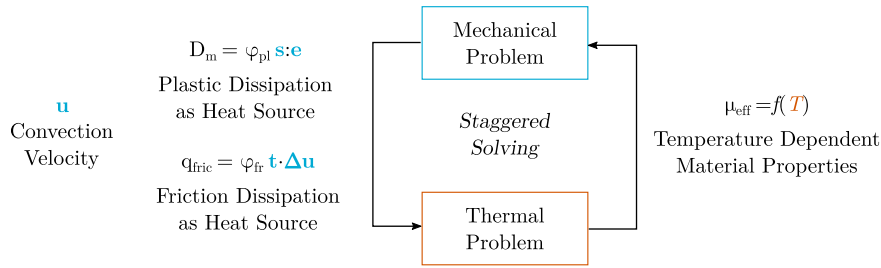


Fig. 4. Thermo-mechanical coupling.

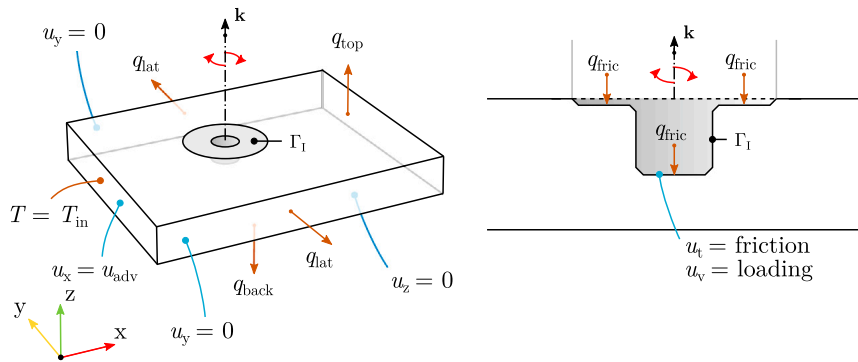


Fig. 5. Thermal and mechanical boundary conditions on the exterior boundaries (left) and the tool-workpiece interface Γ_1 , visualized as cross section of domain (right).

Thermal. The strong form of the thermal problem reads

$$\rho_0 c \left(\frac{1}{\alpha} \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) + \nabla \cdot \mathbf{q} = D_m \quad \text{with} \quad \mathbf{q} = -k \nabla T, \quad (3)$$

where the temperature T is the primary unknown, $\mathbf{u}$ the velocity (responsible for the convective heat transfer), ρ_0 the material density, c the specific heat capacity, k the thermal conductivity and $\mathbf{q}$ the heat flux. D_m denotes the fraction of the plastic dissipation that is turned into heat (see Section 2.3). Note that $c(T)$ and $k(T)$ are temperature-dependent material parameters. The parameter α is introduced to artificially accelerate the transient phase by increasing the thermal diffusion. Thus, the steady state condition can be achieved faster. This method was first introduced by Dialami et al. [23].

2.2. Boundary conditions

Fig. 5 shows the thermal and mechanical boundary conditions of the problem. The mechanical friction and loading conditions, as well as the thermal friction dissipation condition are applied on the tool-workpiece interface $\Gamma_1 \subset \Gamma$, a subset of the domain's boundary.

Mechanical loading. The mechanical problem of Eqs. (1) and (2) is complemented by the boundary conditions

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{on} \quad \Gamma_u \quad \text{and} \quad (4)$$

$$\mathbf{s}\mathbf{n} + p\mathbf{n} = \bar{\mathbf{t}} \quad \text{on} \quad \Gamma_t, \quad (5)$$

where $\bar{\mathbf{u}}$ and $\bar{\mathbf{t}}$ are the prescribed velocities and tractions, respectively. The boundary conditions on $\Gamma \cap \Gamma_1$ (excluding the interface), are shown in Fig. 5.

The vertical loading/no penetration condition restricts the velocity component *normal* to Γ_1 . For this purpose the normal projector is introduced as $P_n = \mathbf{n} \otimes \mathbf{n}$. On surfaces parallel to the pin axis ($\mathbf{n} \cdot \mathbf{k} = 0$), the velocity is restricted to 0 (no penetration) and on surfaces perpendicular to the pin axis the velocity is restricted to a given velocity $\bar{u}_v$, to account for the vertical tool loading. On surfaces that are neither parallel nor perpendicular to the pin axis, the velocity is constrained to a proportion of u_v . Hence,

$$\mathbf{u} \cdot P_n = -\bar{u}_v \mathbf{k} \cdot P_n. \quad (6)$$

Remark 1. There are two different approaches to account for the vertical loading: the application of a uniform vertical velocity field (Dirichlet conditions) or a uniform pressure distribution (Neumann's condition). Since the tool is modeled as a rigid body, the Dirichlet boundary conditions approach is chosen. See [9].

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

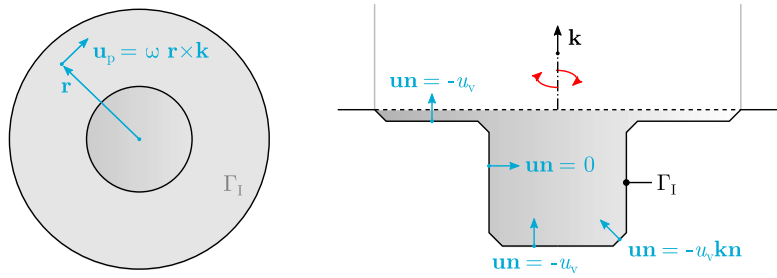


Fig. 6. Kinematics and exterior normals on tool-workpiece interface Γ_I : top view (left), and cross section (right).

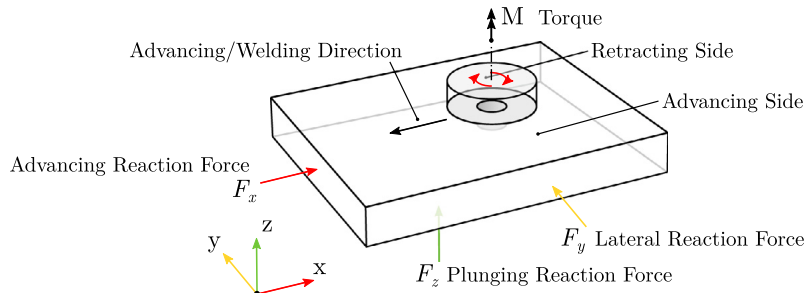


Fig. 7. Reaction forces and torque definitions.

Mechanical friction. Friction only occurs *in plane* with Γ_I , hence the friction condition affects all velocity components *except* the ones normal to the surface. For this purpose the tangential projector is introduced as $P_t = \mathbf{1} - P_n$. The applied tractions are

$$\bar{\mathbf{t}} = \bar{\mathbf{t}}_f = a \underbrace{\|(\bar{\mathbf{u}}_p - \mathbf{u}) \cdot P_t\|^{q-1}}_{a^*} (\bar{\mathbf{u}}_p - \mathbf{u}) \cdot P_t, \quad (7)$$

where $\bar{\mathbf{u}}_p$ is the pin velocity at the interface, $\mathbf{u}$ is the material velocity at the interface (see Fig. 6), a is the friction parameter and q is the sensitivity parameter. The term a^* determines how strongly $(\bar{\mathbf{u}}_p - \mathbf{u}) \cdot P_t \approx 0$ is enforced.

Dialami et al. [9] and Dialami et al. [1] showed that good results can be achieved when a Modified Newton Friction Law is used. Here the friction parameter a is modified geometrically to account for the non-uniform pressure field under the tool shoulder. For moderate values of a some amount of slip occurs, whereas for large values a *full stick* condition is recovered (i.e., Regular Newton Friction Law).

$$a(x, y, \psi) = \frac{1}{2} \left(a_{\max} + a_{\min} + (a_{\max} - a_{\min}) \tanh \frac{6(x \cos \psi + y \sin \psi)}{R} \right) \quad (8)$$

The friction parameter a depends on the position relative to the pin rotation axis (distances x , y and angle ψ), as shown in Eq. (8). The function yields high friction values ($a \rightarrow a_{\max}$) in front of the tool and low friction values ($a \rightarrow a_{\min}$) behind the tool. Dialami et al. [24] show that by introducing the in-plane rotation angle ψ , the obtained reaction forces (Fig. 7) are in even better agreement with experimental results.

Fig. 8 shows plots of Eq. (8). On the left, a is depicted as a function of only x , with $\psi = 0^\circ$. On the right, a is shown as a function of both x and y with an in-plane angle $\psi > 0^\circ$. Note that the slope of the transition from $a_{\min}$ to $a_{\max}$ also depends on the radius R . For the part of Γ_I that is the pin (depicted in blue), the radius of the pin is used $R = R_{\text{pin}}$. For the remaining part of Γ_I that is the shoulder (depicted in orange), the radius of the shoulder is used $R = R_{\text{shd}}$.

The effect of the different parameters is shown qualitatively in Fig. 9. Fig. 9a shows that with a full stick model ($a_{\min} = a_{\max}$) the resulting tractions are of equal magnitude. In terms of the resulting reaction forces their effect cancels out, both in welding direction and perpendicular to it. Previous works have shown that this leads to a substantial underestimation of the lateral reaction forces F_y . By introducing an $a_{\min}$ value that is low enough to permit slipping in the rear part of the tool, the tractions in $-y$ decrease, which yields a net reaction force in $+y$, as shown in Fig. 9b. Introducing the in-plane rotation angle ψ allows for even greater control. As the angle is increased, the lateral net force remains largely unchanged (especially for small angles), while the net reaction force in welding direction decreased as the forces in x no longer cancel out.

A study of the influence of the different friction parameters is presented in Section 4.2.

Thermal. The thermal problem of Eq. (22) is complemented by the boundary conditions

$$T = \bar{T} \quad \text{on} \quad \Gamma_T, \quad \text{and} \quad (9)$$

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

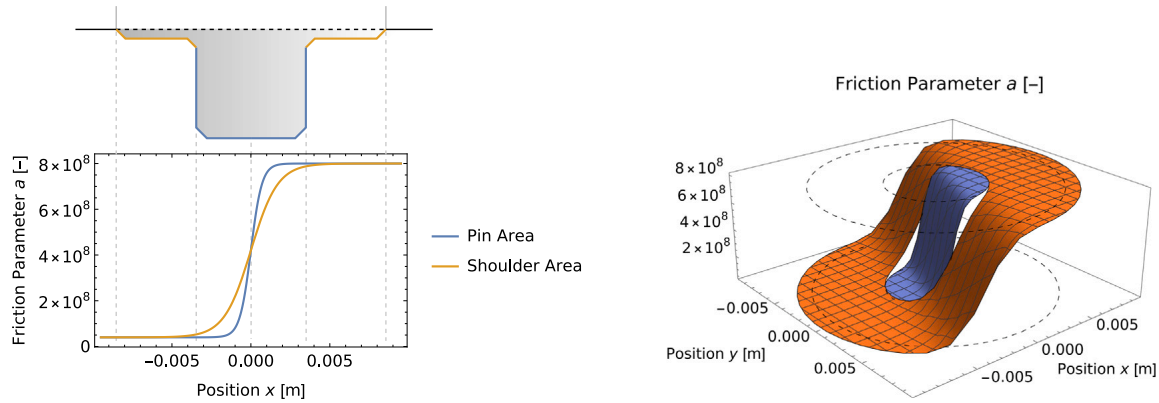


Fig. 8. Geometrically modified friction parameter $a(x, y, \psi)$. Pin radius R_{pin} at 0.0035 m, shoulder radius R_{shd} at 0.0085 m (dotted lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

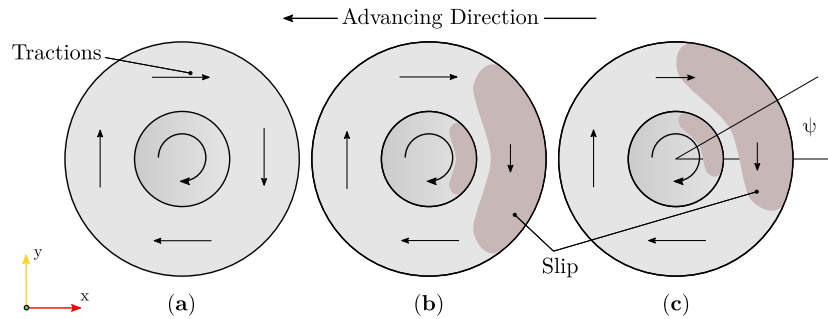


Fig. 9. Different slip patterns for (a) Regular Newton (Full Stick), (b) Modified Newton with $\psi = 0$, and (c) Modified Newton with $\psi > 0$. Coloring: $\square \mathbf{u} \approx \mathbf{u}_p$ (material velocity close to pin velocity, i.e., sticking) and $\blacksquare \mathbf{u} < \mathbf{u}_p$ (material velocity lower than pin velocity, i.e., slipping).

$$\mathbf{q} \cdot \mathbf{n} = \bar{q} \quad \text{on} \quad \Gamma_q, \quad (10)$$

where $\bar{T}$ are the prescribed temperatures and $\bar{q}$ are the prescribed heat fluxes. The boundary conditions on $\Gamma \cap \Gamma_I$ (excluding the interface), are shown in Fig. 5. The heat fluxes $\bar{q}_{\text{lat}}$, $\bar{q}_{\text{top}}$, $\bar{q}_{\text{back}}$ are modeled as Robin boundary conditions, as shown in Eq. (11), where $h(T)$ is the heat transfer coefficient and T the temperature at domain boundary and T_{env} the environment temperature.

$$\bar{q} = h(T)(T - \bar{T}_{\text{env}}) \quad (11)$$

The heat fluxes on Γ_I , caused by the friction dissipation are computed as

$$\bar{q}_{\text{fric}} = \varphi_{\text{fr}} \bar{\mathbf{t}}_f \cdot (\bar{\mathbf{u}}_p - \mathbf{u}) \cdot \mathbf{P}_t, \quad (12)$$

based on the tangential components of the tractions and velocity differences, as introduced previously. The proportionality factor φ_{fr} determines how much friction dissipation is converted into heat. Based on the thermal conductivity properties of the tool and workpiece materials, it can be determined how much heat is absorbed by either part. See [3] for a more detailed explanation. In this work it is assumed that 75% of the friction dissipation is converted to heat and half of it is absorbed by the workpiece, resulting in $\varphi_{\text{fr}} = 0.375$.

2.3. Material modeling

Visco-plastic behavior is widely used to describe the non-Newtonian material flow of metal-forming processes, including FSW. The overall behavior is *shear-thinning*, which means that under low strain rates, the material response is very viscous, while under high strain rates, the material becomes less viscous. In this work, the material behavior is described using the Norton–Hoff law [25,26], which models the strain rate-stress relation as a power law

$$\mathbf{s} = 2\mu_{\text{eff}} \mathbf{e} = 2\mu_{\text{eff}} \nabla^s \mathbf{u} \quad \text{with} \quad \mu_{\text{eff}}(T, \gamma) = \mu_0(T) \gamma^{m(T)-1}, \quad (13)$$

where $\mathbf{e} = \nabla^s \mathbf{u}$ is the deviatoric strain rate, $\nabla^s \mathbf{u}$ being the symmetric gradient of the velocity, $\gamma = \sqrt{2}\|\mathbf{e}\|$ the equivalent strain rate, $\tau = \mu_{\text{eff}} \gamma$ the equivalent stress, and μ_0 the consistency parameter. Finally, m is the so-called flow index. Shear-thinning behavior is observed for $m < 1$, and Newtonian behavior is recovered for $m = 1$.

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

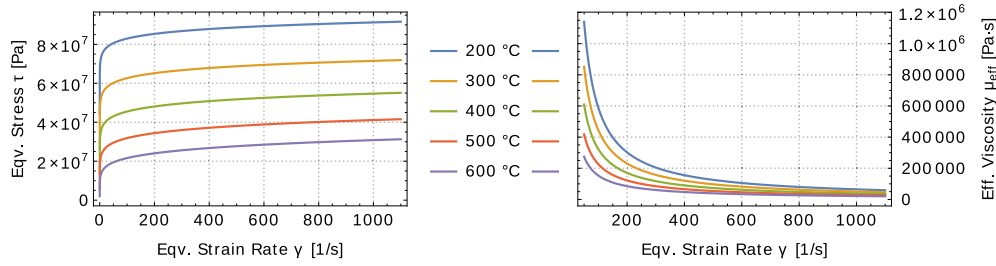


Fig. 10. Temperature and strain rate dependent effective viscosity and equivalent stress.

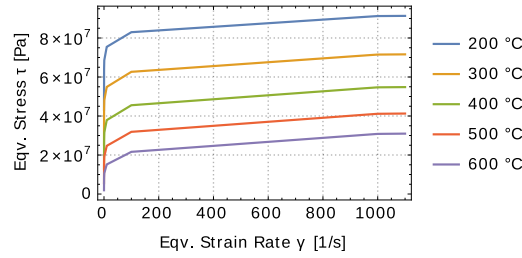


Fig. 11. Linearized material law.

The plastic dissipation is computed as $\mathbf{s} : \mathbf{e}$, where the operator, $:$ is defined as the double contraction of two tensors (i.e., $s_{ij}e_{ij}$), and the proportionality factor φ_{pl} determines how much of the plastic dissipation is accounted for as a heat source for the thermal problem. In this work a value of $\varphi_{pl} = 0.75$ is assumed.

$$D_m = \varphi_{pl} \mathbf{s} : \mathbf{e} \quad (14)$$

The model's temperature dependency further influences its non-linear behavior. Higher temperatures lead to lower viscosities, and a weaker shear thinning effect, as illustrated in Fig. 10.

Due to the highly non-linear nature of the presented material law, the authors have proposed a linearized version. Dialami et al. [23], Venghaus et al. [27]. The power law is converted to a piecewise linear function by evaluating the original law at strain rates of 1.0, 10.0, 100.0, and 1000.0 1/s and linearly interpolating the results, as seen in Fig. 11. The linearized material law performs superior with regard to convergence rates and numerical stability.

2.4. Finite element approximation

Mechanical. Mixed finite element formulations are only naturally stable (i.e., they satisfy the *inf-sup* condition, [28]) for certain combinations of shape functions. To use arbitrary shape function orders (including linear/linear) stabilization techniques need to be applied. In this work the Orthogonal Subgrid Scale stabilization (OSS) is used. See [29] and [30] for an in depth derivation of the OSS stabilization and [22] for its application to FSW problems. Additional stabilization terms are introduced in Section 3.2, which are linked to EDM.

The stabilized weak form of the problem stated in Eqs. (1) and (2) reads

$$\begin{aligned} \int_{\Omega} \nabla^s \delta \mathbf{u}_h \cdot \mu_{eff} \nabla^s \mathbf{u}_h dV + \int_{\Omega} (\nabla \cdot \delta \mathbf{u}_h) p_h dV + \int_{\Omega} \delta p_h (\nabla \cdot \mathbf{u}_h) dV \\ - \underbrace{\sum_K \tau_K \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] dV}_{OSS} = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} dV + \int_{\Gamma_t} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} dA, \end{aligned} \quad (15)$$

The unknowns are $\mathbf{u}_h = \mathbf{N}_u \mathbf{U}_h$ and $p_h = \mathbf{N}_p \mathbf{P}_h$, where $\mathbf{U}_h$ and $\mathbf{P}_h$ are the nodal values of the velocity and pressure fields. $\mathbf{N}_u$ and $\mathbf{N}_p$ are the corresponding shape functions. The test functions are defined using the same set of shape functions (Galerkin method), hence $\delta \mathbf{u}_h = \mathbf{N}_u \delta \mathbf{U}_h$ and $\delta p_h = \mathbf{N}_p \delta \mathbf{P}_h$.

The OSS stabilization parameter is

$$\tau_K = c_1 \frac{h^2}{\mu_{eff}}, \quad (16)$$

where h is the element size and c_1 is the (user defined) stabilization constant. The operator $\mathbf{P}_h^\perp$ denotes the projection onto the space orthogonal to the finite element space.

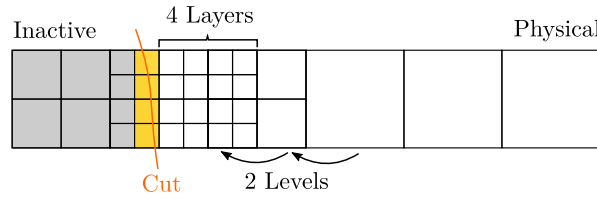


Fig. 12. Refinement levels and layers.

Thermal. Convection dominated thermal problems can experience instability. In this work, Variational Multi-Scale (VMS) stabilization is used to address this issue. See [31] for a comparison of different stabilization techniques for diffusion–convection–reaction equations (including VMS) and [22] on the stabilization of FSW problems.

Note that the terms containing second order derivatives (i.e., the Laplacian operators Δ), disappear when using linear elements.

The stabilized weak form of Eq. (3) reads

$$\begin{aligned} \int_{\Omega} \delta T_h \rho_0 c \left(\frac{1}{\alpha} \frac{\partial T_h}{\partial t} + \mathbf{u}_h \cdot \nabla T_h \right) dV + \int_{\Omega} \nabla \delta T_h \cdot k \nabla T_h dV \\ + \underbrace{\sum_K \varsigma_K \int_{K(\Omega)} \mathbf{u}_h \nabla \delta T_h \cdot \mathbf{P}_h^T [\mathbf{u}_h \nabla T_h] dV}_{\text{OSS}} = \int_{\Omega} \delta T_h D_m dV - \int_{\Gamma_q} \delta T_h \bar{q} dA, \end{aligned} \quad (17)$$

with

$$\varsigma_K = \left(c_2 \frac{k}{h^2} + c_3 \frac{\rho_0 c \|\mathbf{u}\|}{h} \right)^{-1}. \quad (18)$$

The unknown is $T_h = \mathbf{N}_T \mathbf{T}_h$, where $\mathbf{T}_h$ are the nodal values of the temperature and $\mathbf{N}_T$ are the corresponding shape functions. The tests functions are defined using the same shape functions (Galerkin method), hence $\delta T_h = \mathbf{N}_T \delta \mathbf{T}_h$ and c_2 and c_3 are the (user defined) stabilization constants. A BDF1 time integration scheme is used for time discretization.

3. Non-body-fitted FSW simulation

This section outlines the specific EDM used in the present work. The computational domain is split into a physical subdomain (the workpiece), where the governing equations of the thermo-mechanical problem are solved, and a second subdomain (the tool), where a harmonic extension of the physical solution is computed. The subdomains are decoupled, thus, the tool subdomain does not influence the solution in the physical subdomain. This approach was chosen because the tool is considered a rigid body, and heat fluxes between the workpiece and tool are negligible. Therefore, neither a mechanical nor a thermal solution is required in the tool subdomain. Though, the harmonic extensions is necessary to ensure good initial conditions for nodes that change state and become part of the physical subdomain as the tool rotates.

As it is common for EDMs, the interface between physical and inactive subdomains does not align with the mesh; thus, *cut elements* appear. Section 3.1 explains how the embedded tool geometry is made accessible to the finite element algorithm to determine physical, inactive and cut elements.

The present work uses linear tetrahedral elements to discretize the computational domain. Although they are less accurate than higher-order elements or different topologies (e.g., hexahedral), the meshes are easy to generate, and the numerical integration is cheaper due to the lower number of integration points. Most important, the integration of the cut elements is relatively straightforward, as the linear elements can easily be subdivided into boundary conforming sub-integration cells (see Section 3.2).

Finally, linear elements are a good choice when dealing with VMS-stabilized mixed formulations and Ghost Stabilization (see Section 3.2). One reason is that projections to the finite element space (which both methods heavily rely on) can be performed using lumped mass matrices. Since the shape functions are positive throughout the entire element, row summing can be utilized without the risk of negative mass matrix entries.

A sufficient degree of accuracy is achieved by refining the finite element mesh. The refinement criterion of choice is a *Layers-and-Levels* criterion, which ensures a specified number of subdivisions of cut elements (levels) and a specified number of smallest elements next to the interface (layers, see Fig. 12). An additional condition, called two-one-balancing, enforces that the refinement level of neighboring elements can only differ by one.

3.1. Domain description

EDM simulations require classifying elements; in this work the options are physical, inactive or cut. In order to carry out this classification, the finite element algorithm needs access to a description of the shape of the embedded domain (or its boundary). Two viable methods exist: explicit *Boundary Representations* (BREPs) or implicit level set functions. In the following, both approaches are briefly outlined, and the author's choice is motivated.

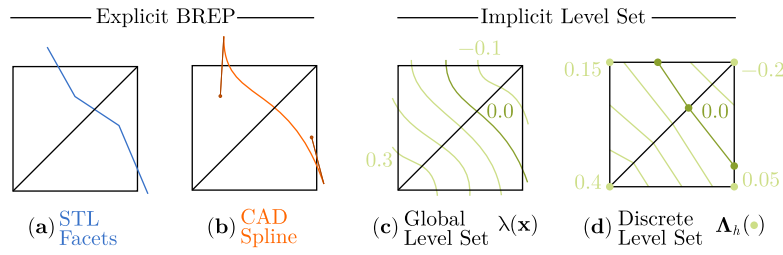


Fig. 13. Selection of domain description methods.

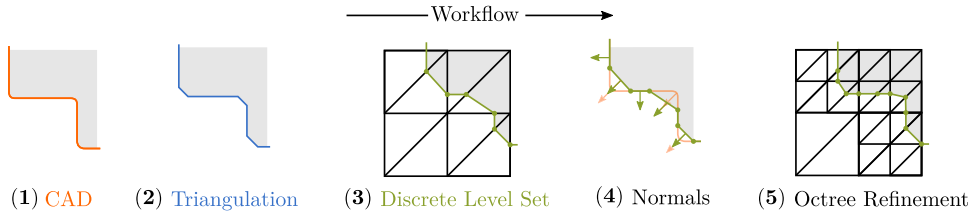


Fig. 14. CAD to simulation workflow.

On the one hand, the use of the BREP approach gives access to an *explicit* description of the embedded domain's shape. This could be facilitated by giving the finite element algorithm access to an STL, STEP, or IGES geometry (see Fig. 13a,b) and performing checks for each element against the BREP to see if (and how) they intersect (see [32]). This approach can yield exceptionally accurate results because it uses the unaltered shape of the embedded domain. However, this approach can be tedious to implement as the resulting physical and inactive sub-integration cells can be as geometrically complex as the original BREP.

On the other hand, Implicit domain descriptions do not provide an explicit description of a shape but easily answer the question of where a given point lies in relation to the described shape. This is commonly known as a level set function [33,34]. An isosurface of this level set function becomes an implicit description of the domain's boundary. A signed distance function is the simplest example of a level set function, where the level set value is the shortest orthogonal distance from a given point to the domain boundary. Positive values describe points outside the domain, negative ones are points inside the domain, and zero values indicate points belonging to the boundary. Level set function can be global, thus defined on the entire domain (Fig. 13c) or discrete and thus are defined for a finite set of points in space (Fig. 13d).

This work uses a discrete level set function. The discrete values Λ_h are obtained by computing the signed distances of the finite element nodes to the surface of a CAD geometry. The values are then interpolated within the elements using shape functions N_λ . The same shape functions are used as for the interpolation of unknowns and spatial coordinates (isoparametric concept).

$$\lambda_h = N_\lambda \Lambda_h \quad (19)$$

This approach is less accurate than the spline-based BREP approach since it relies on the shape functions. However, it is much more robust since it only requires a finite set of level set values, and the resulting physical and inactive sub-integration cells have relatively simple topologies. The full workflow is depicted in Fig. 14 and includes the following steps:

1. Creation of FSW tool geometry in CAD software of choice (STEP/IGES),
2. Conversion into a triangulation (see Remark 2),
3. Computation of discrete signed distance function and interpolation of level set function using element shape functions,
4. Computation of surface normals (see. Remark 3),
5. Octree-based mesh refinement, using layers and levels refinement criteria.

The library used to access the CAD geometry is [35] and the efficient evaluation of the signed distance function is based on works by Slyadnev et al. [36].

Remark 2. Converting the smooth and parameterized CAD geometry into a discrete triangulation seems counter-intuitive since the CAD surface provides a more accurate description of the geometry. However, computing the signed distance function can be computationally demanding, especially if the CAD geometry consists of many irregular spline surfaces. Using a triangulation (similar to an STL surface description) rather than the original parameterized CAD geometry speeds up the computation of the signed distance function. The main advantage is that an octree search algorithm quickly identifies the relevant subset of facets for each FE node. The subsequent computation of the orthogonal distance is performed much faster as the facets of the triangulation are of much simpler topology.

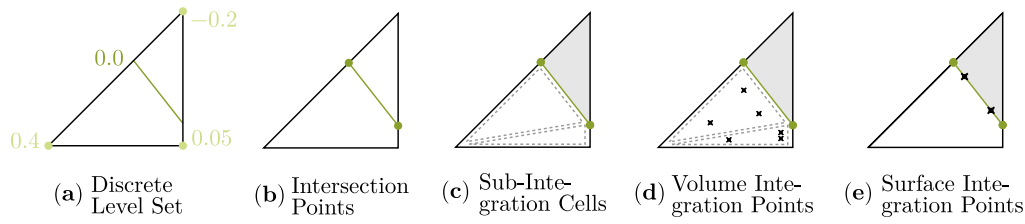


Fig. 15. Domain description methods.

Remark 3. Accurate surface normals are crucial for the correct application of the friction boundary condition (see Section 2.2). Tests by the authors have shown that the normals obtained from the discrete level set function (although they are true to the cut elements) were not suitable and polluted the solution. As a remedy, the surface integration points of the cut elements are projected onto the triangulation of the CAD surface, and a more accurate normal is recovered from the surface itself. Even better normals could be recovered from the parameterized CAD surface, but, as explained in the previous remark, working directly with the CAD surface leads to increased computing times.

3.2. Cut element integration

A major challenge when using EDMs is the integration of cut elements. This includes both the integration of the (discontinuous) volume as well as the integration of the immersed boundary in order to apply boundary conditions. In the end, two sets of integration points (volume and boundary) are required. Different methods can be applied to determine the correct amount of integration points, their location as well as the corresponding weights.

For the present work the boundary conforming subintegration approach is used. Based on the discrete level set function (Fig. 15a), the intersection points are computed on the element boundaries (Fig. 15b). Next, each resulting sub-element is subdivided into sub-integration cells of known topology (e.g., triangles in Fig. 15c, tetrahedra in 3D) and finally Gauss–Legendre integration points are placed in the sub-integration cells. Their weights are adjusted accordingly.

The benefits of this approach are its low computation cost and the relatively low number of resulting integration points (see Figs. 15d and e). Other known approaches are *Moment Fitting* [37,38], which requires solving a least squares problem for each element (computationally expensive but very versatile), and octree subintegration, which is very stable but results in a much larger number of integration points. However, there are methods to reduce the final number of integration points in octree subintegration by merging sub cells [39].

Octree subintegration is frequently used for higher order elements since the integration accuracy can be easily increased by adding few more refinement levels. This is not possible for the boundary conforming subintegration approach presented in this work. However, when using linear elements and a discrete level set approach, the geometrical approximation error is dominant due to the use of linear shape functions to interpolate the nodal level set values. Consequentially, the boundary conforming subintegration approach can accurately integrate the cut elements with a minimal number of integration points. To achieve more accurate results the geometrical approximation error is tackled by simply refining the mesh, as shown in Fig. 14 (5).

Dirichlet boundary conditions. Since Γ_1 is an immersed boundary, Dirichlet boundary conditions cannot be enforced in strong form by prescribing the nodal unknowns. Therefore, the vertical loading/no penetration condition is applied using the penalty method and is integrated on the right hand side of Eq. (15). The prescribed tractions are thus expanded as $\tilde{\mathbf{t}} = \tilde{\mathbf{t}}_f + \tilde{\mathbf{t}}_v$ with

$$\tilde{\mathbf{t}}_v = b(-\bar{u}_v \mathbf{k} - \mathbf{u}) \cdot \mathbf{P}_n. \quad (20)$$

The penalty factor to enforce the velocity constraint is b and it is chosen large enough to ensure $(-\bar{u}_v \mathbf{k} - \mathbf{u}) \cdot \mathbf{P}_n \approx 0$.

Small cut stabilization. Independently from the employed integration approach, small cut elements pose a problem. They considerably worsen the conditioning of the linear system, which is detrimental when using iterative solvers. The present work uses the *Ghost Stabilization* approach by Burman [40] to overcome the problem. The conditioning is enhanced by stabilizing the gradients of the unknowns (i.e., $\nabla \mathbf{u}$ and ∇p), using orthogonal projections to the finite element space. In contrast to OSS stabilization, Ghost stabilization is only applied to the cut elements, and to the first two layers of physical elements $K(\Omega^{1st2})$, as shown in Fig. 16. The same applies for the orthogonal projections $\mathbf{P}_h$ required for Ghost Stabilization.

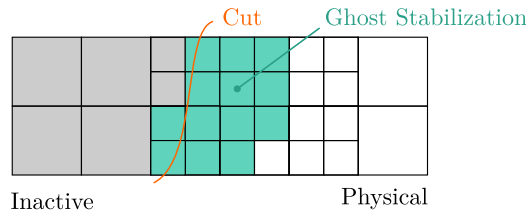


Fig. 16. Ghost Stabilization on cut elements and first two physical layers ■.

The fully stabilized weak form of the mechanical problem reads

$$\begin{aligned} \int_{\Omega} \nabla^s \delta \mathbf{u}_h \cdot \mu_{\text{eff}} \nabla^s \mathbf{u}_h \, dV + \int_{\Omega} (\nabla \cdot \delta \mathbf{u}_h) p_h \, dV + \int_{\Omega} \delta p_h (\nabla \cdot \mathbf{u}_h) \, dV \\ - \underbrace{\sum_K \tau_K \int_{K(\Omega)} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h] \, dV}_{\text{OSS}} \\ + \sum_K \gamma_{1K} \int_{K(\Omega^{1\text{st}})} \nabla \delta \mathbf{u}_h \cdot \mathbf{P}_h^\perp [\nabla \mathbf{u}_h] \, dV \\ + \underbrace{\sum_K \gamma_{2K} \int_{K(\Omega^{1\text{st}})} \nabla \delta p_h \cdot \mathbf{P}_h^\perp [\nabla p_h - \mathbf{b}] \, dV}_{\text{GS}} = \int_{\Omega} \delta \mathbf{u}_h \cdot \mathbf{b} \, dV + \int_{\Gamma_t} \delta \mathbf{u}_h \cdot \bar{\mathbf{t}} \, dA, \end{aligned} \quad (21)$$

with the stabilization parameters

$$\gamma_{1K} = c_4 \frac{h^2}{\tau_K} \quad \text{and} \quad \gamma_{2K} = c_5 \tau_K, \quad (22)$$

where c_4 and c_5 are (user defined) stabilization constants. Note that τ_K stems from the definition of the Variational Multiscale Stabilization (see Eq. (16)).

The temperature problem is stabilized in an equivalent way, penalizing the temperature gradients ∇T . The fully stabilized weak form reads

$$\begin{aligned} \int_{\Omega} \delta T_h \rho_0 c \left(\frac{1}{\alpha} \frac{\partial T_h}{\partial t} + \mathbf{u}_h \cdot \nabla T_h \right) \, dV + \int_{\Omega} \nabla \delta T_h \cdot k \nabla T_h \, dV \\ + \underbrace{\sum_K \varsigma_K \int_{K(\Omega)} \mathbf{u} \nabla \delta T_h \cdot \mathbf{P}_h^\perp [\mathbf{u} \nabla T_h] \, dV}_{\text{OSS}} \\ + \underbrace{\sum_K \epsilon_K \int_{K(\Omega^{1\text{st}})} \nabla \delta T \cdot \mathbf{P}_h^\perp [\nabla T] \, dV}_{\text{GS}} = \int_{\Omega} \delta T_h D_m \, dV - \int_{\Gamma_q} \delta T_h \bar{q} \, dA, \end{aligned} \quad (23)$$

with the stabilization parameter

$$\epsilon_K = c_6 \frac{h^2}{\varsigma_K}, \quad (24)$$

where c_6 is a (user defined) stabilization constant. Note that ς_K stems from the definition of the VMS stabilization (see Eq. (18)).

3.3. Harmonic extensions

The computational domain is divided into the workpieces (physical subdomain) and the tool. Assuming a rigid tool and neglecting the heat fluxes between tool and workpiece, the subdomain representing the pin tool is inactive. Hence, nothing is assembled in the corresponding elements and the inactive subdomain is therefore completely disregarded and does not influence the solution of the physical one.

However, problems occur when the level set function is updated every time step, as it is the case for non-rotationally symmetric tools. Consequentially, some FE nodes change from being part of the inactive subdomain to the physical one. The lack of realistic initial conditions for the new physical nodes harms convergence. As a solution, the inactive subdomain is not entirely disregarded but a harmonic extension is computed instead (see [21]). This is done for both the thermal and the mechanical problem and ensures good initial conditions for the non-linear iterative process.

Given the relative simplicity of the thermal problem, the same convection–diffusion equation is also solved in the inactive subdomain. Only the two subdomains are uncoupled by eliminating all contributions from inactive elements to nodes that belong to cut elements during assembly.

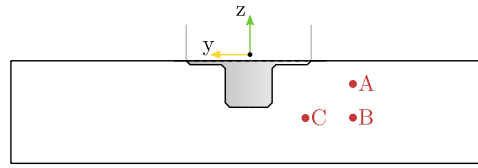


Fig. 17. Position of thermocouples in relation to tool.

Table 1

Thermocouple coordinates.

	y [mm]	z [mm]
A	-11.0	-2.0
B	-11.0	-5.0
C	-5.0	-5.0

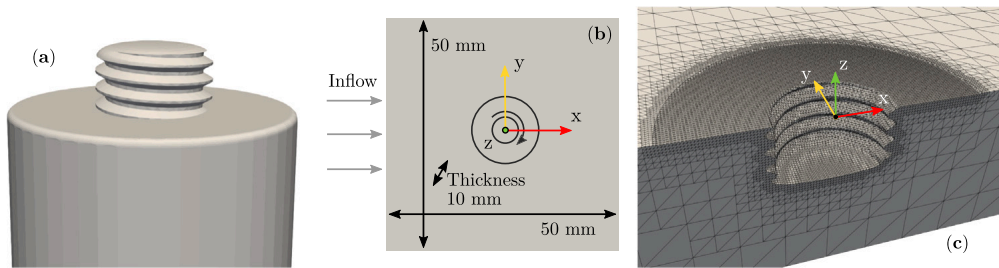


Fig. 18. (a) Tool geometry, (b) computational domain and its dimensions, and (c) cross section of mesh.

However, a similar approach for the mechanical problem failed since the very non-linear, temperature-dependent Stokes flow is a much more complicated problem, and numerical instabilities were observed when also solving the Stokes flow equations in the inactive subdomain. As a remedy, a much simpler, one-dimensional Poisson problem is solved for each of the four unknowns: u_x , u_y , u_z , and p . Again, both subdomains are uncoupled.

4. Numerical examples

The numerical examples are selected to show how the EDM approach can be used to simulate FSW problems of industrial interest and achieve a good agreement between experimental data and simulation results. First, the parameter fitting process is explained, using available experimental data as reference. Then, the numerical results are analyzed and the sensitivity analysis of different processing parameters is presented.

To assess the quality of the results, the reaction forces and torque (F_x , F_y , F_z , M , see Fig. 7), as well as temperature evolution at three positions (T_{A-C}) are considered. The temperature evolution was recorded as functions of time $T(t)$, using thermocouples (K -type, 0.5 mm) at fixed (material) positions, shown in Fig. 17 and Table 1, respectively. Since the simulation is carried out in a Eulerian framework, the experimental *temperature-over-time* functions were converted to the frame of reference of the simulation, becoming *temperature-over-distance* functions $T(x)$. This is possible because of the uniform flow at the thermocouple positions, where u_y and $u_z \approx 0$.

Fig. 18a shows the tool geometry, the computational domain is of size 50 mm $\times$ 50 mm $\times$ 10 mm (Fig. 18b), and a detail of the mesh (Fig. 18c). A summary of all process and numerical parameters is given in Table 2. The temperature-dependent material parameters of the used aluminum alloy (Al6063-T6) μ_0 , m , c and k are shown in Figs. 20 and 21. Fig. 19 shows a cross section of the computational domain, highlighting the different element types (physical, inactive and cut) as well as the zero iso surface of the level set function.

The computational domain consists of approximately 115 000 nodes and 458 000 tetrahedral elements. On average 300 000 elements (66%) are classified as physical, 83 000 (18%) as cut and 75 000 (16%) as inactive. Each simulation was carried out for 0.3 s, with a time step size of 0.0025 s, which corresponds to 120 steps. Given a tool rotation frequency of $\omega = 600$ RPM (10 Hz), this leads to a tool rotation of 9° per step and 40 time steps per revolution. After about 40 time steps (1 full revolution) a periodic steady-state solution is obtained. The reaction forces presented in Sections 4.2, 4.3 and 4.5 are averaged over the last two full rotations (steps 41 to 120). Distributed memory parallelization using MPI [41] is employed. Each simulation took about 2 h 40', using 8 processors at a time.

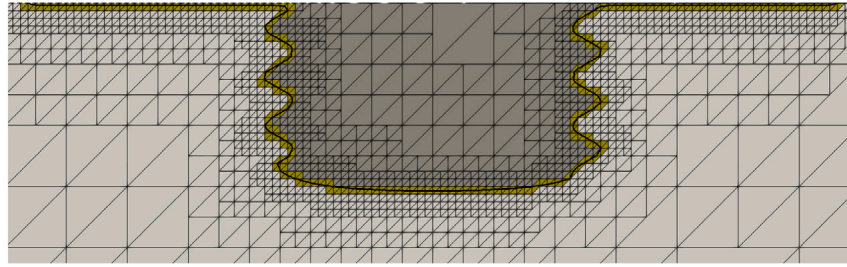


Fig. 19. Level set iso line and element types: Physical ■, Cut ■ and Inactive ■.

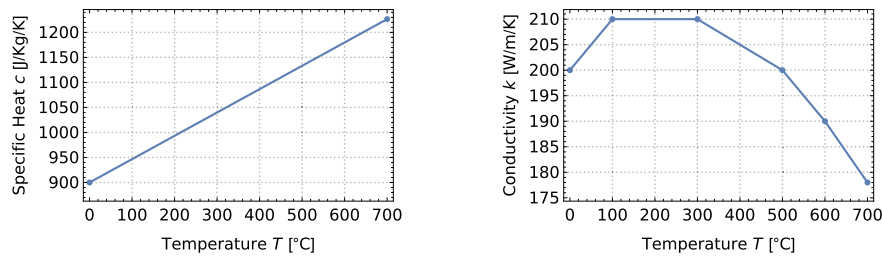


Fig. 20. Temperature dependent specific heat capacity and thermal conductivity of Al6063-T6.

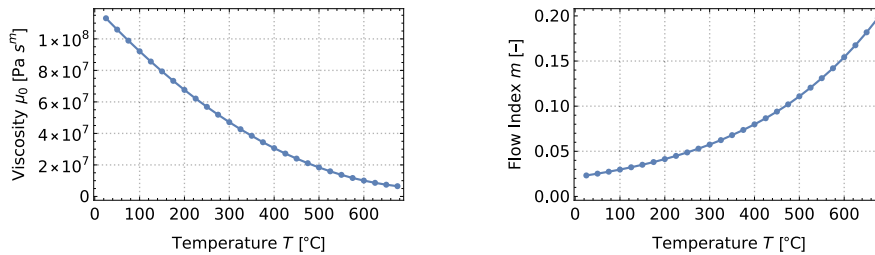


Fig. 21. Temperature dependent Norton-Hoff Consistency and Flow Index of Al6063-T6.

4.1. Mesh convergence study

Convergence studies for immersed methods, using the method of manufactured solutions (MoMS) have been presented by Burman et al. [42] and Petö et al. [43], among others. Regarding the specific problem, presented in this work, a mesh convergence study was conducted to observe the behavior of the reaction forces F_x , F_y , and F_z , the torque M , and the maximum temperature $T_{\max}$ for different levels of mesh refinement (see Fig. 12).

For the coarsest mesh, three levels of mesh refinement were applied, resulting in approximately 160 000 elements. For the finest mesh, seven levels were applied, resulting in approximately 50 Million elements. In order to encourage a smooth solution, an unfeatured, cylindrical tool and a *full-stick* friction law ($a_{\min} = a_{\max} = 420 \cdot 10^6$) were selected.

The results are shown in Fig. 22, indicating convergence for all variables. The convergence rates for the reaction forces are approximately 1.5, while the convergence rate for the torque is approximately 2.0. Both values are above 1.0, which is the expected value for stresses using linear shape functions and, therefore, for reaction forces as well.

4.2. Parameter study: Friction, loading and dissipation

The phenomenological friction law (as introduced in Section 2.2) is mainly controlled by three parameters: the penalty parameters $a_{\max}$, $a_{\min}$ and the *in-plane* rotation angle ψ . For the parameter fitting, $a_{\max}$ is fixed at $800 \cdot 10^6$ as this ensures a sticking condition in front of the tool. $a_{\min}$ is varied between $45 \cdot 10^6$ and $62 \cdot 10^6$ to allow for different levels of slip on the back part of the tool. The *in-plane* rotation angle ψ is varied between 12.25° and 16.25° .

Fig. 23 shows that modifying $a_{\min}$ and ψ affects the reaction forces F_x and F_y , while F_z , the torque M and the temperatures remain unchanged.

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

Table 2
Process and numerical parameters.

Parameter		Value	Unit	Parameter		Value	Unit
Advancing Velocity	u_{adv}	400	mm/min	Inflow Env. Temp.	$\bar{T}_{in}$	37	°C
Tool Rotation Frequency	ω	600	RPM	Backplate Env. Temp.	$\bar{T}_{back}$	20	°C
Tool Axis	$\mathbf{k}$	$\mathbf{e}_z$	–	Lateral Env. Temp.	$\bar{T}_{lat}$	50	°C
Plunging Depth	d	0.05	mm	Top Env. Temp.	$\bar{T}_{top}$	20	°C
				Backplate HTC	h_{back}	2500	W/(m ² K)
Vertical Loading	u_v	2.96	mm	Lateral HTC	h_{lat}	50000	W/(m ² K)
Upper Friction Penalty	a_{max}	$800 \cdot 10^6$	–	Top HTC	h_{top}	100	W/(m ² K)
Lower Friction Penalty	a_{min}	$58 \cdot 10^6$	–	Specific Heat Capacity	c	Fig. 20	J/Kg/K
Sensitivity Parameter	q	0.1	–	Thermal Conductivity	k	Fig. 20	W/m/K
Loading/No Pen. Penalty	b	$5.0 \cdot 10^{11}$	–	Density	ρ_0	2700	kg/m ³
In-Plane Angle	ψ	14.25	° (Deg)	Thermocouple Positions		Fig. 17	
Transition Slope	s	4.0	–				
				Shoulder Radius	R_{sh}	8.5	mm
Refinement Levels		5	–	Pin Radius	R_{pin}	3.5	mm
Refinement Layers		1	–	Pin Length	l_{pin}	3.8	mm
Stabilization Constants	c_1	1.0	–	Thread Type		M7	Left Hnd.
	c_2	4.0	–	Speed-Up	α	100	–
	c_3	2.0	–				
	c_4	1.0	–	Aluminum Alloy		Al6063-T6	
	c_5	5.0	–	Norton-Hoff Consistency	μ_0	Fig. 21	Pa s ^m
	c_6	0.001	–	Norton-Hoff Flow Index	m	Fig. 21	
				Material Dissipation Fra.	ϕ_{pl}	0.75	–
Element Size (finest)	h	~ 0.16	mm	Friction Dissipation Fra.	ϕ_{fr}	0.375	–
Number Nodes		~ 115 000	–				
Number Elements		~ 458 000	–				

Table 3
Calibrated friction, loading and dissipation parameters.

Parameter	a_{min}	ψ	u_v	ϕ_{pl}	ϕ_{fr}
	$58.0 \cdot 10^6$	14.25°	2.96 mm/s	0.75	0.375

The vertical velocity u_v (Fig. 24) is varied between 2.56 mm/s and 3.36 mm/s, being related to the vertical reaction force F_z , while leaving the other reactions unchanged.

The material dissipation fraction ϕ_{pl} (Fig. 25) is varied between 0.68 and 0.82. This corresponds to converting 68 to 82% of plastic dissipation into heat. Note that the friction dissipation fraction ϕ_{fr} is pegged to the plastic dissipation fraction in a ratio of $\phi_{fr} = \phi_{pl}/2$. The trend clearly shows that the higher the fractions ϕ_{pl} and ϕ_{fr} are, the higher the maximum temperature and, in turn, the lower the reaction forces and torque. This relationship can be easily explained by the softening of the material at higher temperatures.

Fig. 26 shows the temperature curves for observation point A (see Fig. 17). Here, the same trend as for the maximum temperature can be observed. While the friction and loading parameters have little to no effect on the temperatures, the dissipation fractions ϕ_{pl} and ϕ_{fr} clearly have.

4.3. Model calibration

The friction law, vertical velocity, and dissipation parameters introduced in the previous section were fitted to obtain a good match with experimental results in terms of the reaction force and temperature evolution. The fitting was carried out using DAKOTA [44]. Table 3 shows the resulting parameters.

The resulting reaction forces and torque are shown in Table 4, showing a good agreement between simulation results and experimental data. Fig. 27 shows the comparison of the three temperature profiles, achieving an exceptional agreement with the experimental data. The experimental data was obtained by our industrial partner NORSK HYDRO ASA and was first published by Dialami et al. [1].

4.4. Numerical results analysis

The velocity contour (Fig. 28a) shows that the highest velocities occur in the boundary layer, especially on the outermost perimeter of the shoulder. It can also be seen that high velocity gradients occur at the shoulder surface, while rapidly decreasing in the interior. However, the side view shows that in the area where the shoulder meets the pin, moderate velocities occur, progressing

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

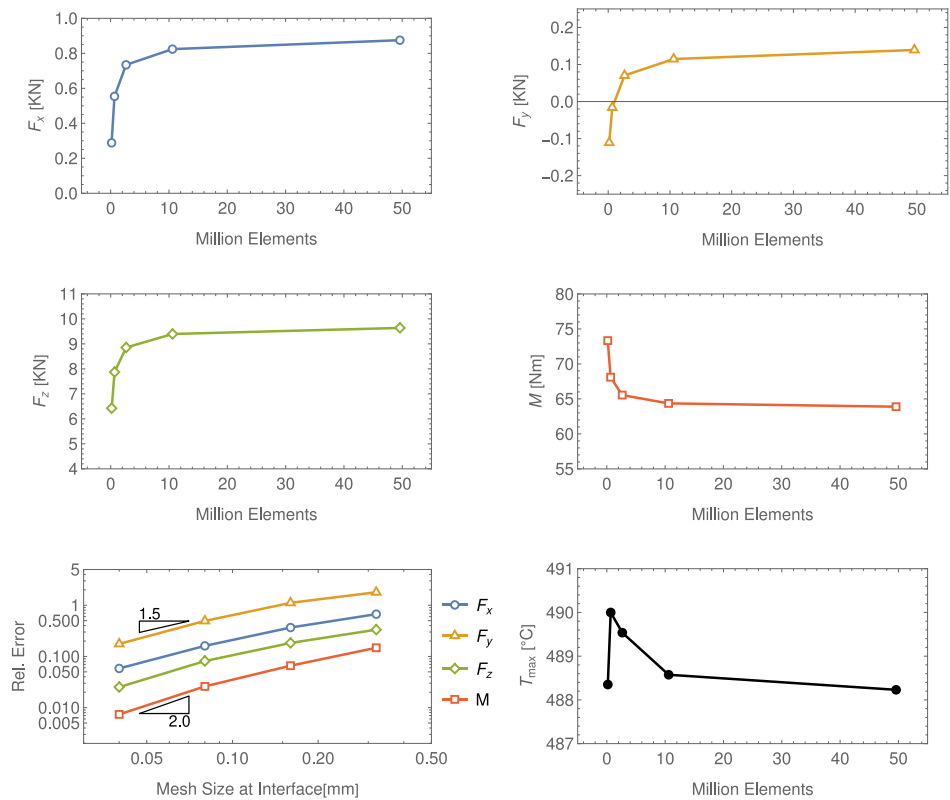


Fig. 22. Evolution of reaction forces, torque and maximum temperature due to mesh refinement, as well as convergence rate of reaction forces and torque (bottom left).

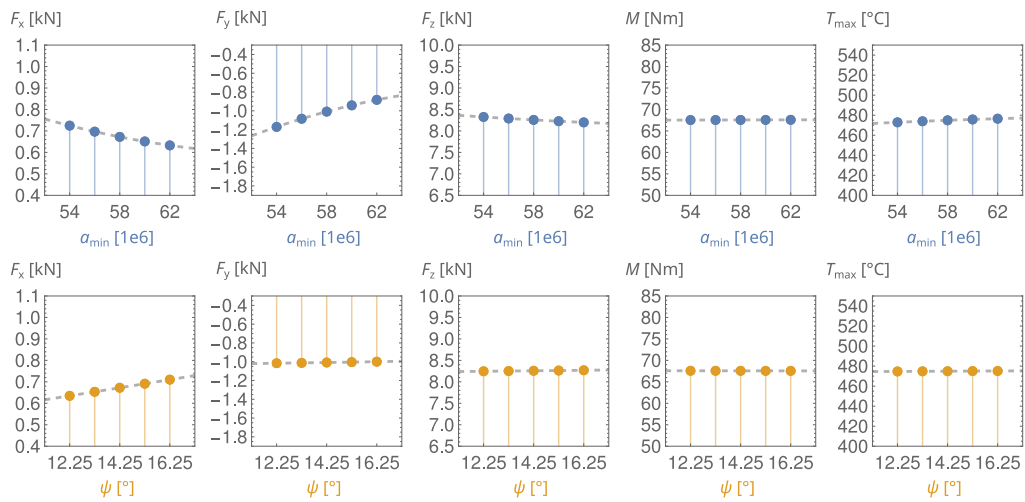


Fig. 23. Reaction forces, torque and maximum temperature in response to friction parameters a_{min} and ψ .

Table 4
Reaction forces and torque, experiments vs. simulations.

	F_x	F_y	F_z	M
Experimental Data	700.0 N	-1000.0 N	8000.0 N	62.0 Nm
Fitting	672.3 N	-1008.5 N	8257.1 N	67.5 Nm
Error	4.0%	0.9%	3.2%	8.9%

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

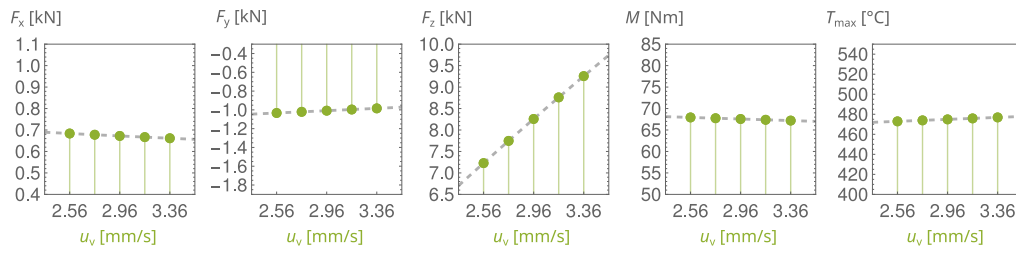


Fig. 24. Reaction forces, torque and maximum temperature in response to friction parameter u_v .

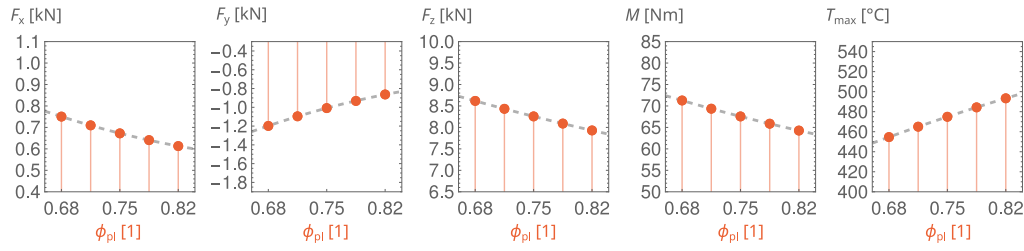


Fig. 25. Reaction forces, torque and maximum temperature in response to friction parameter ϕ_{pl} .

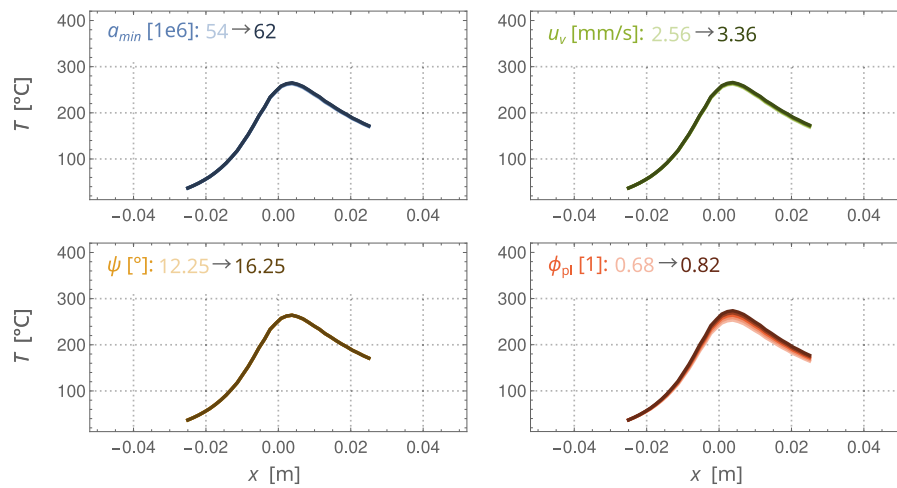


Fig. 26. Temperature at point A in response to parameter changes.

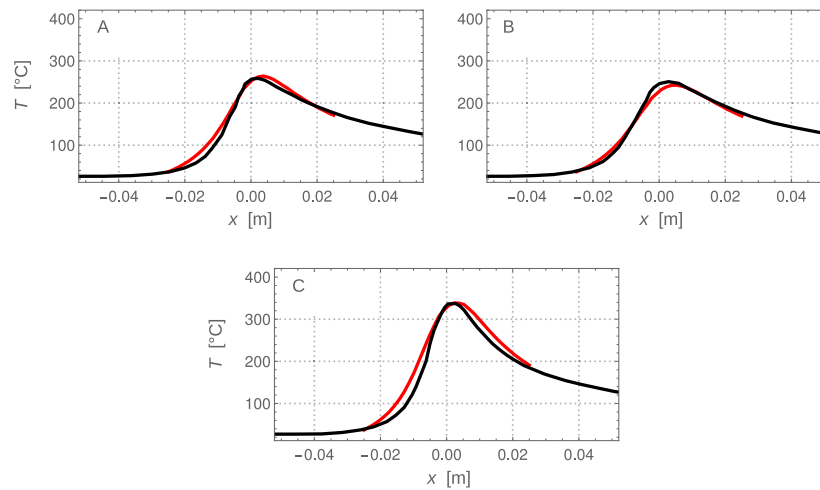


Fig. 27. Experimental temperatures ■ vs. temperatures obtained from simulation ■.

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

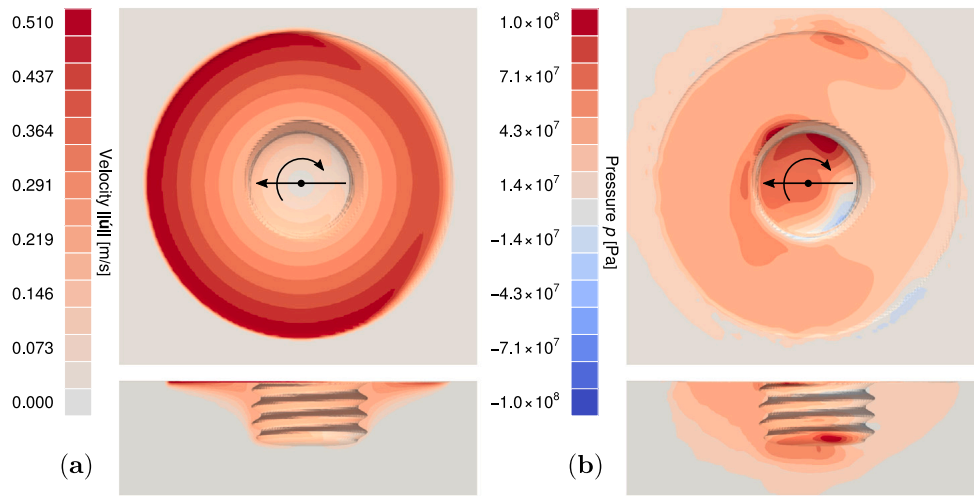


Fig. 28. Contour plots of velocity $\mathbf{u}$ and pressure p .

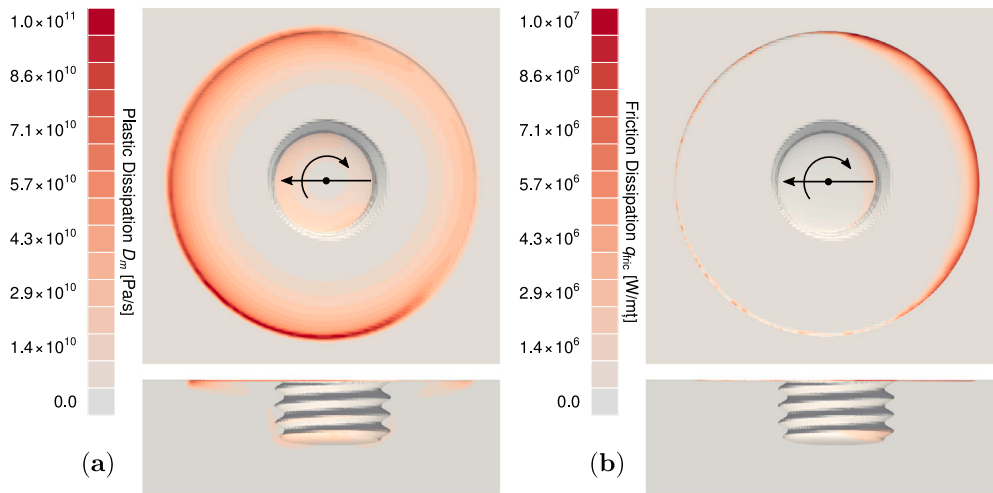


Fig. 29. Contour plots of plastic dissipation D_m and friction dissipation $\bar{q}_{fric}$.

deeper into the material. The transition from full stick on the left to moderate slip on the right and the in-plane rotation angle ψ are clearly visible in the top view.

The pressure field (Fig. 28b) is reasonably smooth, showing the highest values where the friction condition transitions from slipping to sticking and vice versa. The field is dominated by the pressure buildup due to the vertical loading and the advancing of the tool.

Figs. 29a and b depict the plastic and friction dissipation, respectively. Naturally, the plastic dissipation is highest where the largest strain rates occur, and the friction dissipation is highest where the most slip occurs (see velocity field).

The temperature field (Fig. 30a) shows a consistent inflow temperature of 37 °C. The highest temperatures is located in the contact area, and the effect of the heat convection due to the tool movement is also evident.

Fig. 30b shows the effective viscosity μ_{eff} , being highly dependent on the temperature field and the strain rate. Low viscosities are caused by high temperatures (softening) and high strain rates (shear-thinning nature of the Norton-Hoff law). As a result of both effects, the lowest viscosities occur in the stir zone.

4.5. Sensitivity analysis: Processing parameters and tool size

This study aims to examine the influence of processing parameters on reaction forces and heat generation. Fig. 31 shows that increasing the rotational speed leads to a decrease in all reaction forces and an increase in the temperature field (see also temperature evolution in Fig. 35, top left). The decrease in reaction forces is due to increased temperatures which are responsible for the material softening.

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

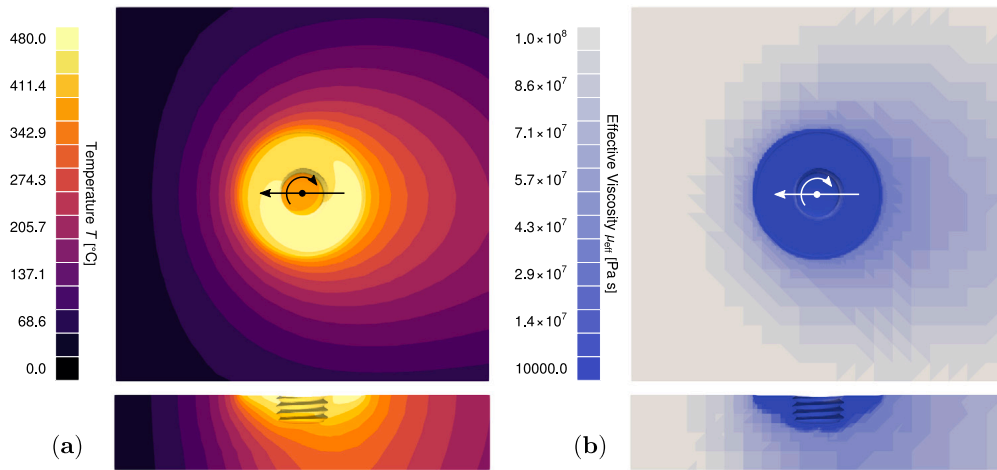


Fig. 30. Contour plots of temperature T and effective viscosity μ_{eff} .

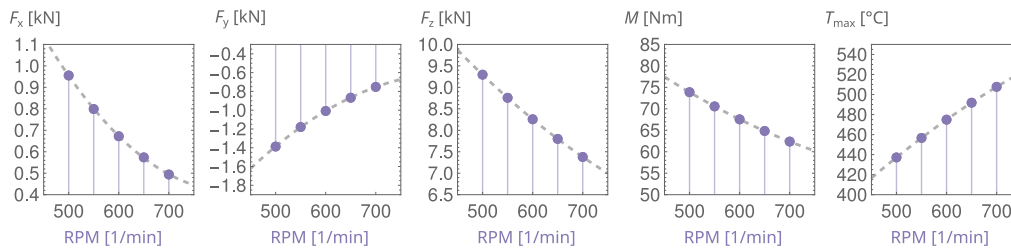


Fig. 31. Reaction forces, torque and maximum temperature in response to rotational velocity ω .

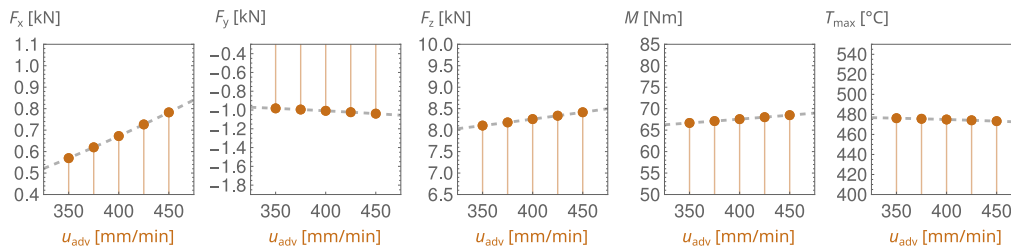


Fig. 32. Reaction forces, torque and maximum temperature in response to advancing velocity u_{adv} .

The tool advancing velocity u_{adv} affects the longitudinal reaction force F_x and leaves all other reaction forces relatively unchanged (see Fig. 32). This behavior is very much in line with the actual FSW process experience. Note that the maximum temperature marginally decreases as the tool advancing velocity is increased. The more pronounced heat convection explains this effect.

The plunging depth d affects the advancing force F_x , the torque M , as well as the maximum temperature T_{max} and the temperature evolution $T(x)$, as shown in Fig. 33. The deeper the tool is plunged into the material, the larger the contact area between the tool and the workpiece is, leading to larger forces in welding direction, torque and maximum temperature. The temperature increase recorded in Fig. 35 has most likely two causes: the generally increased heat generation due to the larger contact area and the closer proximity to the thermocouple.

For the last test (see Fig. 34), the tool has been scaled to sizes ranging from 90 and 110% of its original dimensions. On the one hand, the longitudinal force F_x and the lateral force F_y show a similar behavior, compared to the first test (Fig. 31), where the rotational speed has been changed. The cause is most likely the increased temperatures, which are similar to the first test (rotational speed). As a result, F_x and F_y decrease for larger tools. However, the decrease of F_x is less substantial as in the first case, although the increase in temperature is comparable.

On the other hand, the torque M and plunging force F_z show a different trend. This is caused by the larger cross section of the scaled-up tools. The increased shoulder radius causes the increase in torque — dominating the softening of the material due to the increased temperatures. The increased shoulder radius is also responsible for the increased plunging forces F_z .

H. Venghaus et al.

Computer Methods in Applied Mechanics and Engineering 435 (2025) 117539

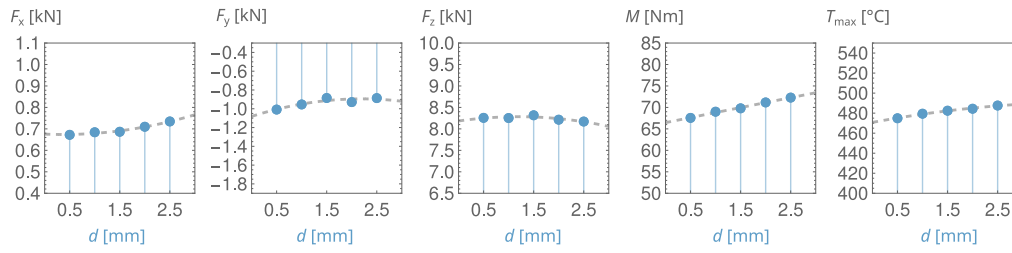


Fig. 33. Reaction forces, torque and maximum temperature in response to plunging depth d .

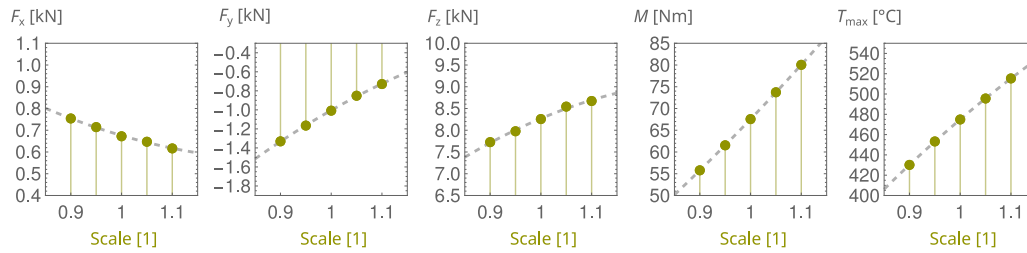


Fig. 34. Reaction forces, torque and maximum temperature in response to the tool scale.

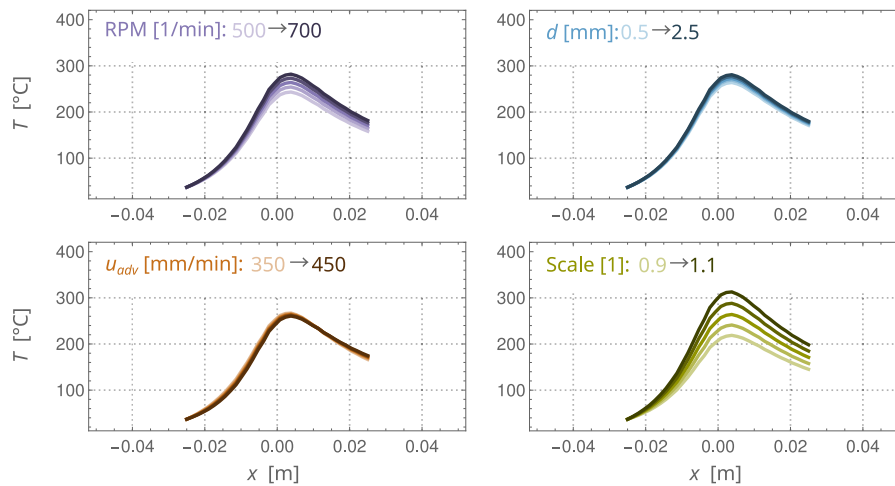


Fig. 35. Temperature at point A in response to parameter changes.

Note that the tool scale has a bigger influence on the temperature evolution (Fig. 35, bottom right) than the rotational speed (Fig. 35, top left). At the same time, the maximum temperatures T_{max} of both tests (Figs. 31 and 34) show very similar values. This effect can be explained with the changing distances between tool and thermocouple in the scaled tool analysis.

This study clearly shows that the developed model allows for an accurate analysis of the FSW process. Parameter changes cause a clear trend in the thermo-mechanical behavior of the FSW process in line with expectations. Through the optimization of process parameters it is possible to achieve desired outcomes in terms of reaction forces and temperature, thereby providing control over the process.

For instance, higher advancing speeds are needed to speed up the overall welding process. The model predicts increased advancing reaction forces F_x as a result, but the rotational velocity can now be adjusted by the correct amount to counter that effect.

Similarly, to weld thicker plates, a larger tool might be chosen to ensure that the stir zone protrudes deeper into the plates. Again, the rotational velocity can be adjusted accordingly to ensure that the peak temperatures remain constant even though a larger tool is used.

5. Conclusions

This work presents an Embedded Finite Element approach to simulate Friction Stir Welding processes and overcome the limitations of (purely) Lagrangian, Eulerian, or Arbitrary Lagrangian Eulerian approaches. The thermo-mechanical problem is solved

by, splitting the computational domain in a non-mesh conforming manner into a physical part (workpiece) and an inactive one (pin-tool). The resulting cut elements are integrated using a boundary conforming subintegration approach. The domain decomposition is accomplished using the level set function, which is directly computed from CAD geometry. Numerical stability is addressed by stabilizing the mechanical mixed velocity–pressure problem and the convection dominated thermal problem using the Variational Multiscale Method, by computing a harmonic extension in the inactive subdomain, and by employing the Ghost Stabilization technique to stabilize small cuts.

The friction law is calibrated by adjusting a set of numerical friction parameters and comparing the resulting reaction forces and temperature evolution curves to experimental data. The results are in excellent agreement with the experimental data, showing the validity of the new approach.

Furthermore, a sensitivity analysis of different processing parameters such as advancing speed, rotational speed, and tool size, is presented, showing that the model yields results that align with the expectations of FSW processes. Hence, the model can be used to adjust processing parameters in physical FSW applications in a targeted manner, allowing for greater control over the process.

The authors look forward to using the developed model to analyze the effect of tool tiling, various tool features such as complex shoulder geometries, shoulder-pin aspect ratios and threading as well as the prediction of welding defects.

CRedit authorship contribution statement

H. Venghaus: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **M. Chiumenti:** Writing – review & editing, Supervision, Methodology, Conceptualization. **J. Baiges:** Writing – review & editing, Validation, Supervision, Software. **D. Juhre:** Writing – review & editing, Supervision. **N. Dialami:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

H. Venghaus acknowledges the support by CIMNE, a Severo Ochoa Centre of Excellence (2019–2023) under the grant [CEX2018-000797-S](#), funded by the Research Agency of the Spanish State (AEI, [10.13039/501100011033](#)) and the support from the PRiMuS Project (Printing pattern based and MultiScale enhanced performance analysis of advanced Additive Manufacturing components, [PID2020-115575RB-I00](#)) through the Research Agency of the Spanish State (AEI, [10.13039/501100011033](#)). The authors further acknowledge the valuable expertise in the field of Friction Stir Welding and the provision of experimental data from the industrial partner NORSK HYDRO ASA.

Data availability

Data will be made available on request.

References

- [1] Narges Dialami, et al., Experimental validation of an FSW model with an enhanced friction law: Application to a threaded cylindrical pin tool, *Metals* (ISSN: 2075-4701) 7 (11) (2017) 491, <http://dx.doi.org/10.3390/met7110491>.
- [2] M. Chiumenti, et al., Numerical modeling of friction stir welding processes, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 254 (2013) 353–369, <http://dx.doi.org/10.1016/j.cma.2012.09.013>.
- [3] N. Dialami, et al., Challenges in thermo-mechanical analysis of friction stir welding processes, *Arch. Comput. Methods Eng.* 24 (1) (2016) 189–225, <http://dx.doi.org/10.1007/s11831-015-9163-y>.
- [4] R. Padmanaban, V. Ratna Kishore, V. Balusamy, Numerical simulation of temperature distribution and material flow during friction stir welding of dissimilar aluminum alloys, *Procedia Eng.* 97 (2014) 854–863, <http://dx.doi.org/10.1016/j.proeng.2014.12.360>.
- [5] Saad B. Aziz, et al., Impact of friction stir welding (FSW) process parameters on thermal modeling and heat generation of aluminum alloy joints, *Acta Metall. Sin. (Eng. Lett.)* 29 (9) (2016) 869–883, <http://dx.doi.org/10.1007/s40195-016-0466-2>.
- [6] Rahul Kumar, V. Pancholi, R.P. Bharti, Material flow visualization and determination of strain rate during friction stir welding, *J. Mater. Process. Technol.* 255 (2018) 470–476, <http://dx.doi.org/10.1016/j.jmatprotec.2017.12.034>.
- [7] Narges Dialami, Miguel Cervera, Michele Chiumenti, Defect formation and material flow in Friction Stir Welding, *Eur. J. Mech. A Solids* 80 (2020) 103912, <http://dx.doi.org/10.1016/j.euromechsol.2019.103912>.
- [8] C. Leitão, et al., Material flow in Friction Stir Welding, *Microsc. Microanal.* 14 (S3) (2008) 87–90, <http://dx.doi.org/10.1017/s1431927608089472>.
- [9] N. Dialami, et al., Enhanced friction model for Friction Stir Welding (FSW) analysis: Simulation and experimental validation, *Int. J. Mech. Sci.* 133 (2017) 555–567, <http://dx.doi.org/10.1016/j.ijmecsci.2017.09.022>.
- [10] D.G. Andrade, et al., Analysis of contact conditions and its influence on strain rate and temperature in friction stir welding, *Int. J. Mech. Sci.* 191 (2021) 106095, <http://dx.doi.org/10.1016/j.ijmecsci.2020.106095>.
- [11] Philippe Bussetta, et al., Comparison of a fluid and a solid approach for the numerical simulation of friction stir welding with a non-cylindrical pin, *Steel Res. Int.* 85 (6) (2013) 968–979, <http://dx.doi.org/10.1002/srin.201300182>.
- [12] Narges Dialami, et al., An apropos kinematic framework for the numerical modeling of friction stir welding, *Comput. Struct.* (ISSN: 0045-7949) 117 (2013) 48–57, <http://dx.doi.org/10.1016/j.compstruc.2012.12.006>.

- [13] N. Sukumar, et al., Extended finite element method for three-dimensional crack modelling, *Internat. J. Numer. Methods Engrg.* (ISSN: 1097-0207) 48 (11) (2000) 1549–1570, [http://dx.doi.org/10.1002/1097-0207\(20000820\)48:11<1549::aid-nme955>3.0.co;2-a](http://dx.doi.org/10.1002/1097-0207(20000820)48:11<1549::aid-nme955>3.0.co;2-a).
- [14] Thomas-Peter Fries, Ted Belytschko, The extended/generalized finite element method: An overview of the method and its applications, *Internat. J. Numer. Methods Engrg.* (ISSN: 1097-0207) 84 (3) (2010) 253–304, <http://dx.doi.org/10.1002/nme.2914>.
- [15] Jamshid Parvizian, Alexander Düster, Ernst Rank, Finite cell method, *Comput. Mech.* 41 (1) (2007) 121–133, <http://dx.doi.org/10.1007/s00466-007-0173-y>.
- [16] A. Düster, et al., The finite cell method for three-dimensional problems of solid mechanics, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 197 (45–48) (2008) 3768–3782, <http://dx.doi.org/10.1016/j.cma.2008.02.036>.
- [17] A.M. Roma, C.S. Peskin, M.J. Berger, An adaptive version of the immersed boundary method, *J. Comput. Phys.* 153 (2) (1999) <http://dx.doi.org/10.1006/jcph.1999.6293>.
- [18] Isabelle Ramière, Philippe Angot, Michel Belliard, A fictitious domain approach with spread interface for elliptic problems with general boundary conditions, *Comput. Methods Appl. Mech. Engrg.* 196 (4–6) (2007) 766–781, <http://dx.doi.org/10.1016/j.cma.2006.05.012>.
- [19] Anita Hansbo, Peter Hansbo, An unfitted finite element method, based on Nitsche's method, for elliptic interface problems, *Comput. Methods Appl. Mech. Engrg.* 191 (47–48) (2002) 5537–5552, [http://dx.doi.org/10.1016/s0045-7825\(02\)00524-8](http://dx.doi.org/10.1016/s0045-7825(02)00524-8).
- [20] Uwe Janoske, Markus Bürger, Tobias Geiger, Modelling of friction stir welding processes with a coupled immersed boundary-/volume of fluid-approach, *Prog. Comput. Fluid Dyn. Int. J.* (ISSN: 1741-5233) 19 (1) (2019) 55, <http://dx.doi.org/10.1504/pcfd.2019.097602>.
- [21] Ramon Codina, Guillaume Houzeaux, Herbert Coppola-Owen, Joan Baiges, The fixed-mesh ALE approach for the numerical approximation of flows in moving domains, *J. Comput. Phys.* (ISSN: 0021-9991) 228 (5) (2009) 1591–1611, <http://dx.doi.org/10.1016/j.jcp.2008.11.004>.
- [22] C. Agelet de Saracibar, et al., Computational modeling and sub-grid scale stabilization of incompressibility and convection in the numerical simulation of friction stir welding processes, *Comput. Methods Engrg.* (ISSN: 1886-1784) 21 (1) (2014) 3–37, <http://dx.doi.org/10.1007/s11831-014-9094-z>.
- [23] N. Dialami, et al., A fast and accurate two-stage strategy to evaluate the effect of the pin tool profile on metal flow, torque and forces in friction stir welding, *Int. J. Mech. Sci.* (ISSN: 0020-7403) 122 (2017) 215–227, <http://dx.doi.org/10.1016/j.ijmecsci.2016.12.016>.
- [24] Narges Dialami, Miguel Cervera, Michele Chiumenti, Effect of the tool tilt angle on the heat generation and the material flow in friction stir welding, *Metals* (ISSN: 2075-4701) 9 (1) (2018) 28, <http://dx.doi.org/10.3390/met9010028>.
- [25] F.H. Norton, *The creep of steel at high temperatures*, Bulletin, McGraw-Hill book Company, Incorporated, 1929.
- [26] Nicholas John Hoff, Approximate analysis of structures in the presence of moderately large creep deformations, *Quart. Appl. Math.* 12 (1954) 49–55.
- [27] H. Venghaus, et al., An accurate approach to simulate friction stir welding processes using adaptive formulation refinement, *Finite Elem. Anal. Des.* (ISSN: 0168-874X) 224 (2023) 103986, <http://dx.doi.org/10.1016/j.finela.2023.103986>.
- [28] Franco Brezzi, Michel Fortin (Eds.), *Mixed and Hybrid Finite Element Methods*, Springer New York, 1991, <http://dx.doi.org/10.1007/978-1-4612-3172-1>.
- [29] Ramon Codina, Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods, *Comput. Methods Appl. Mech. Engrg.* 190 (13–14) (2000) 1579–1599, [http://dx.doi.org/10.1016/s0045-7825\(00\)00254-1](http://dx.doi.org/10.1016/s0045-7825(00)00254-1).
- [30] Ramon Codina, Stabilized finite element approximation of transient incompressible flows using orthogonal subscales, *Comput. Methods Appl. Mech. Engrg.* 191 (39–40) (2002) 4295–4321, [http://dx.doi.org/10.1016/s0045-7825\(02\)00337-7](http://dx.doi.org/10.1016/s0045-7825(02)00337-7).
- [31] Ramon Codina, Comparison of some finite element methods for solving the diffusion-convection-reaction equation, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 156 (1–4) (1998) 185–210, [http://dx.doi.org/10.1016/s0045-7825\(97\)00206-5](http://dx.doi.org/10.1016/s0045-7825(97)00206-5).
- [32] Manuel Meßmer, Tobias Teschemacher, Lukas F. Leidinger, Roland Wüchner, Kai-Uwe Bletzinger, Efficient CAD-integrated isogeometric analysis of trimmed solids, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 400 (2022) 115584, <http://dx.doi.org/10.1016/j.cma.2022.115584>.
- [33] N. Sukumar, D.L. Chopp, N. Moës, T. Belytschko, Modeling holes and inclusions by level sets in the extended finite-element method, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 190 (46–47) (2001) 6183–6200, [http://dx.doi.org/10.1016/s0045-7825\(01\)00215-8](http://dx.doi.org/10.1016/s0045-7825(01)00215-8).
- [34] Frederic Gibou, Ronald Fedkiw, Stanley Osher, A review of level-set methods and some recent applications, *J. Comput. Phys.* (ISSN: 0021-9991) 353 (2018) 82–109, <http://dx.doi.org/10.1016/j.jcp.2017.10.006>.
- [35] OpenCascade, 2024, URL <https://dev.opencascade.org/>.
- [36] Sergey Slyadnev, Alexander Malyshev, Vadim Turlapov, CAD model inspection utility and prototyping framework based on OpenCascade, in: *GraphiCon'17 Proceedings*, 2017, pp. 323–327.
- [37] B. Müller, F. Kummer, M. Oberlack, Highly accurate surface and volume integration on implicit domains by means of moment-fitting, *Internat. J. Numer. Methods Engrg.* (ISSN: 0029-5981) 96 (8) (2013) 512–528, <http://dx.doi.org/10.1002/nme.4569>.
- [38] Simeon Hubrich, et al., Numerical integration of discontinuous functions: moment fitting and smart octree, *Comput. Mech.* (ISSN: 1432-0924) 60 (5) (2017) 863–881, <http://dx.doi.org/10.1007/s00466-017-1441-0>.
- [39] Márton Pető, et al., Octree-based integration scheme with merged sub-cells for the finite cell method: Application to non-linear problems in 3D, *Comput. Methods Appl. Mech. Engrg.* (ISSN: 0045-7825) 401 (2022) 115565, <http://dx.doi.org/10.1016/j.cma.2022.115565>.
- [40] Erik Burman, Ghost penalty, *C. R. Math.* (ISSN: 1778-3569) 348 (21–22) (2010) 1217–1220, <http://dx.doi.org/10.1016/j.crma.2010.10.006>.
- [41] *MPI Forum, A Message-Passing Interface Standard Version 3.1*, Tech. Rep, MPI Forum, 2015.
- [42] Erik Burman, Peter Hansbo, Mats G. Larson, CutFEM based on extended finite element spaces, *Numer. Math.* (ISSN: 0945-3245) 152 (2) (2022) 331–369, <http://dx.doi.org/10.1007/s00211-022-01313-z>.
- [43] Márton Pető, et al., Code verification of immersed boundary techniques using the method of manufactured solutions, *Comput. Mech.* (ISSN: 1432-0924) 73 (6) (2023) 1283–1309, <http://dx.doi.org/10.1007/s00466-023-02411-x>.
- [44] Keith Dalbey, et al., Dakota a multilevel parallel object-oriented framework for design optimization parameter estimation uncertainty quantification and sensitivity analysis: Version 6.12 theory manual, 2020, <http://dx.doi.org/10.2172/1630693>.

