Engineering Fracture Mechanics 266 (2022) 108390

Contents lists available at ScienceDirect

Engineering Fracture Mechanics

journal homepage: www.elsevier.com/locate/engfracmech

Thermodynamically consistent linear-gradient damage model in Abaqus

Gergely Molnár [a,*], Aurélien Doitrand [b], Adrien Jaccon [c,a], Benoit Prabel [c], Anthony Gravouil [a]

[a] Univ Lyon, INSA-Lyon, CNRS UMR5259, LaMCoS, F-69621, France
[b] Univ Lyon, INSA-Lyon, UCBL, CNRS UMR5510, MATEIS, F-69621, France
[c] Université Paris-Saclay, CEA, Service d'Études Mécaniques et Thermiques, 91191, Gif-sur-Yvette, France

ARTICLE INFO

Keywords:
AT1 phase-field model
Abaqus
UMAT
Spectral energy decomposition
Damage gradient

ABSTRACT

Commercial finite element software follow cautiously the numerical methods developed by the scientific community. Even though the phase-field method is not a default option in Abaqus, the code proposes a unique way to implement and test complex, extraneous models.

In the last five years, phase-field fracture models in Abaqus have gained unbelievable popularity among scientists and engineers alike. However, most implementations are based on the quadratic crack representation function, which is easy to solve but has no elastic threshold. Various solutions have been proposed as a workaround to implement the linear damage function, however, none of them obtain consistent results to the original Griffith solution when the length scale is reduced to zero.

This paper presents an energetically consistent linear damage gradient model in Abaqus. The bound constrained optimization is achieved using Lagrange multipliers as an additional degree of freedom.

We show that when the necessary energy corrections are applied, the phase-field simulations are in agreement with the analytical results of linear elastic fracture mechanics. Furthermore, through elaborate benchmark tests, we verified our code and experimentally demonstrated the validity of our implementation.

1. Introduction

Modeling fracture in engineering materials remains a daunting challenge. In the past century, various approaches have been proposed to understand and model crack initiation, propagation, and branching under different mechanical, environmental, and physical conditions. With the improvement of experimental techniques, it has gradually been recognized that simulating a macroscopic fracture is exceptionally challenging. As this phenomenon is an intrinsically multi-scale problem, the models have to consider the effect of even atomic structures in some cases, which of course is impractical from an engineering point of view.

The pioneering researcher to approximate discrete phenomena with a continuous description was van der Waals [1] in 1873. He was the first to describe atomic liquid–gas interfaces with continuous density functions The foundation of the phase-field method was then laid down by Ginzburg and Landau [2] then Cahn and Hilliard [3]. Both works replaced a discrete interface with a continuous function to study superconductivity and diffuse interfaces. Since then, the concept has been used to solve various problems, such as solidification dynamics [4], image processing [5], or even crack propagation [6].

* Corresponding author.
E-mail address: gergely.molnar@insa-lyon.fr (G. Molnár).

https://doi.org/10.1016/j.engfracmech.2022.108390
Received 13 December 2021; Received in revised form 11 February 2022; Accepted 15 March 2022
Available online 26 March 2022
0013-7944/© 2022 Elsevier Ltd. All rights reserved.

In fracture mechanics, B. Bourdin et al. [6] proposed a groundbreaking theory in which the discrete fracture surface [7,8] in the variational formulation of Francfort and Marigo [9] was replaced by a continuous damage density function. Thus, the original minimization problem was transformed using the Mumford and Shah functional [5] (which is one of the limit cases of the Ambrosio and Tortorelli [10] elliptic regularization function).

Fundamentally, Bourdin used the concept of damage mechanics to degrade the material's stiffness and strength. But contrary to Kachanov's work [11], Bourdin calculated the value of the damage (d), not based on stress or strain fields, but using the concept of the phase-field model. By introducing the spatial gradient of the damage variable and an internal length scale (l_c) into the energy equation, the new formulation treated singular defects (*e.g.,* cracks) in a thermodynamically consistent manner. Since the first introduction of the method, several modifications have been proposed, of which one of the most important was the reformulation of the damage density function to accommodate a linear elastic threshold [12]. This step was important because the original quadratic phase-field representation needed a relatively small l_c to simulate engineering problems where elasticity was important.

The first functional Abaqus phase-field implementation to model crack propagation was published by Liu et al. [13]. In their study, the staggered scheme was still solved using an Intel MKL library. A year later, the first publicly available UEL code followed [14]. Since then the Abaqus community has worked hard to implement ductile material models [15–17], cohesive-type fracture models [18–20], stress corrosion [21] and even hydrogen-assisted embrittlement theories [22,23]. The user subroutine options allow the users to change time integration schemes (quasi-Newton [24], BFGS [25]) with a single command. Explicit time integration schemes [26] are slightly more complicated, as the user element has to be reformulated. Fang et al. [15] replaced the mechanical UEL in the original implementation and solved the displacement problem in a UMAT, which significantly sped up the computation. Recently, Navidtehrani et al. [27] proposed a clever technique to use Abaqus's coupled thermal-stress analysis module to solve the phase-field fracture problem in a single UMAT code. There are promising solutions [20,28] in the literature to replace the quadratic phase-field representation with a linear function and an energy threshold. However, a thermodynamically consistent scheme has yet to be proposed in the finite element code Abaqus (see the discussion on existing models).

Linear damage gradient models are predominantly available as in-house codes [12,29–32]. We therefore felt the need for an energetically well-based linear-gradient phase-field model in Abaqus, which was used to graphically demonstrate the different types of spatial discretization errors. Finally, we validated the current implementation and showed its capabilities in static and dynamic 2D and 3D cases through benchmark examples.

The objective was to extend our previous code [14,16] to solve linear damage functions (with an elastic threshold). This paper thus presents a revised finite element structure where the stress problem was solved in a UMAT. The spectral decomposition technique of Bernard et al. [33] was updated to remain stable even in the case when two principal strains are equal. Finally, we describe a failure surface for both linear and quadratic phase-field functions in the space of all three principal stresses to emphasize the microscopic (phenomenological) importance of the internal length scale.

The paper is structured as follows. First, Section 2 introduces the phase-field damage model and the related energy function. Following the theory, the details of the implementation are presented. Section 3 describes the failure envelope and the effect of the spatial discretization errors whereas Section 4 is dedicated to benchmark examples. Finally, after a discussion in Section 5, Section 6 concludes the paper.

2. Methods

This section describes the underlying mathematical and physical description of the phase-field approach in order to model regularized brittle fracture. Progressively, the different energy contributions of the individual phenomena are presented after which the staggered solution of the weakly coupled problems is explained. This is followed by the bound constraint optimization scheme to enforce damage irreversibility. Finally, the most crucial implementation details are given. The energy functional of the complete elasto-dynamic problem involves the following Lagrangian function:

$$\mathcal{L} = D(\dot{\mathbf{u}}) - \Pi(\mathbf{u}, d), \tag{1}$$

where $D(\dot{\mathbf{u}})$ is the kinetic energy:

$$D(\dot{\mathbf{u}}) = \frac{1}{2}\int_{\Omega} \dot{\mathbf{u}}^T \dot{\mathbf{u}} \rho d\Omega, \tag{2}$$

and $\Pi(\mathbf{u}, d)$ is the potential energy:

$$\Pi(\mathbf{u}, d) = E(\mathbf{u}, d) + W(d). \tag{3}$$

In Eq. (2), $\dot{\mathbf{u}}$ contains the nodal velocities, while ρ is the mass density. The potential energy is constructed from the elastic strain energy (E) and the fracture energy (W). All energy components depend on either the phase-field (d) or the displacement field ($\boldsymbol{u}$) and its temporal derivatives.

2.1. Phase-field/gradient-damage models

In contrast to the discrete concept of fracture, the phase-field approach introduces a gradual transition with an auxiliary variable d between intact ($d = 0$) and fully damaged ($d = 1$) materials. The discrete crack surface is regularized by a length scale (l_c), and with the gradient of the damage field, a crack surface density functional is created.

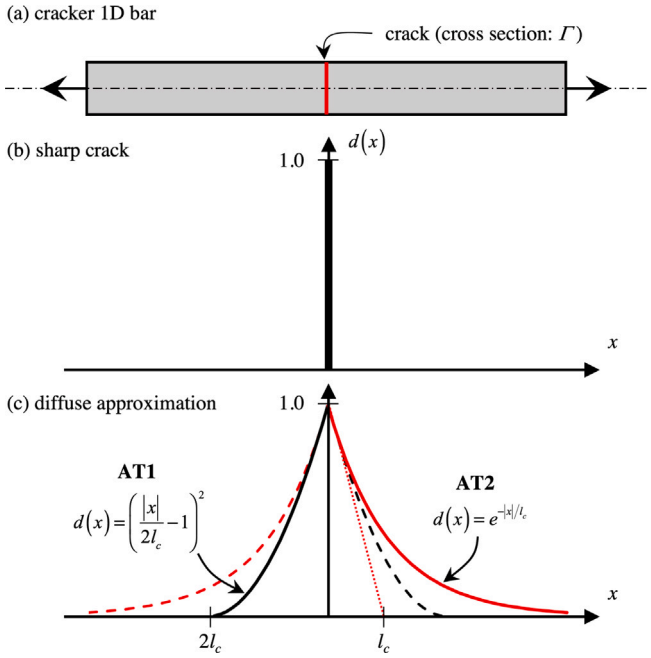

Fig. 1. (a) 1D bar with a crack in the middle and a cross section Γ. (b) Damage field with a sharp crack at $x = 0$. (c) Diffuse representation with models AT1 (black) and AT2 (red) with the length scale parameter l_c.

The total energy dissipated by the opening of the crack can be written as follows:

$$W(d) = \Gamma g_c = \int_\Gamma g_c d\Gamma \approx \int_\Omega g_c \gamma(d, \nabla d)\, d\Omega, \tag{4}$$

where Γ is the size of the discrete crack surface, g_c is Griffith's [7] critical energy release rate, and γ is the crack surface density:

$$\gamma(d, \nabla d) = \frac{1}{c_\omega l_c}\left[\omega(d) + l_c^2 |\nabla d|^2\right]. \tag{5}$$

In the literature, multiple crack representations exist. The so-called geometric function (ω) mostly controls the topology of the phase-field solution. The most widely used models are the AT1 [12] and AT2 [6,34] representations. For the AT1 approach, $\omega(d) = d$, while for AT2, $\omega(d) = d^2$. Wu [30] proposed a mixed representation which renders a cohesive zone model-like behavior.

The total integral on the simulation domain of the crack surface density function should give the theoretical value of the discrete representation. Therefore, depending on the model, γ is normalized by the constant c_ω:

$$c_\omega = 4\int_0^1 \sqrt{\omega(s)}\, ds, \tag{6}$$

with $c_\omega = 8/3$ for AT1 and $c_\omega = 2$ for the AT2 models.

Fig. 1 shows the basic concept of the diffuse damage models with the analytical solutions in 1D for AT1 and AT2 representations.

The advantage of AT1 compared to AT2 is that it has an elastic threshold before failure, whereas for AT2 models the damage appears for any prescribed loading level. On the other hand, AT1 models require a special treatment to enforce positive damage values and damage irreversibility. Further discussion is given in Section 2.4.

2.2. Elastic problem

Damage reduces the elastic strain energy, thus weakening the resistance and the stiffness of the material. The mechanical energy contribution can be written as follows:

$$E(\mathbf{u}, d) = \int_\Omega \psi(\mathbf{u}, d)\, d\Omega, \tag{7}$$

where ψ is the strain energy density. To avoid fracturing under compression ψ is divided into two parts:

$$\psi(\mathbf{u}, d) = g(d)\,\psi_0^+(\varepsilon(\mathbf{u})) + \psi_0^-(\varepsilon(\mathbf{u})), \tag{8}$$

where $g(d)$ is the energy degradation function:

$$g(d) = (1-d)^2 + k, \tag{9}$$

with k being a small number (10^{-12}) responsible for conditioning the solution.

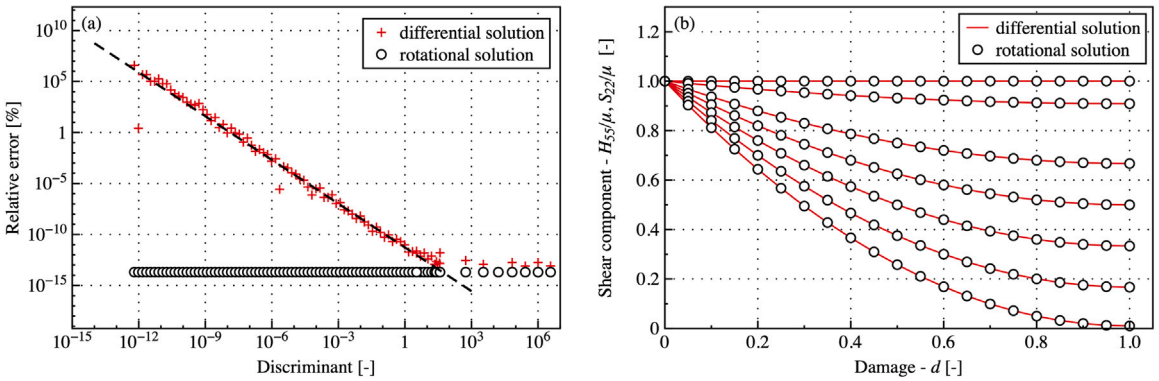

Fig. 2. (a) Maximum relative error between the decomposed and theoretical stiffness matrices ($d = 0$). The error is shown as a function of the discriminant of the third-order polynomial equation to calculate principal strains [16]. (b) Normalized shear component of the stiffness matrix determined using the differential (H_{55}) and the rotational (S_{22}) techniques.

The subscript in Eq. (8) refers to tensile (+) and compressive (−) energies. The degradation function is applied only on the tensile part to avoid crack opening in compression. As a result, compression does not contribute to the crack driving force. Similarly to our latest implementation [16], this study also adopts the 3D spectral decomposition technique [33].

As highlighted previously [16], this technique suffers from a vital numerical flaw, which is usually overcome by introducing a small perturbation: when two eigenvalues out of three are equal, the solution becomes singular. One would think that this occurs only in rare cases, when in fact our most common examples fall in this category, *e.g.,* uniaxial or equibiaxial configurations. Therefore, the present paper describes how we updated the original spectral decomposition technique to give a more robust and faster solution.

In order to stabilize the solution, we first expressed the complete stiffness tensor in the principal directions, and then rotated it in the direction of the original reference system:

$$\mathbf{H} = \mathbf{T}_{\sigma}^{-1}\hat{\mathbf{H}}\mathbf{T}_{\varepsilon}. \tag{10}$$

Here, $\hat{\mathbf{H}}$ is the stiffness matrix in the principal directions, $\mathbf{T}_{\varepsilon}$ and $\mathbf{T}_{\sigma}$ are rotation matrices in Voigt notation compiled from the original basis of the strain tensor. The details are given in Appendix A.

The stiffness matrix in the principal directions can be written as the following block matrix:

$$\hat{\mathbf{H}} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{bmatrix}, \tag{11}$$

where $\mathbf{L}$ is the well-known [33] 3×3 matrix deduced from the potential energy:

$$\mathbf{L} = \frac{\partial^2 \psi}{\partial \hat{\varepsilon}^2} = \frac{E}{1+\nu}\begin{bmatrix} g\left(a_1 \cdot d\right) + g\left(a \cdot d\right)\chi & g\left(a \cdot d\right)\chi & g\left(a \cdot d\right)\chi \\ & g\left(a_2 \cdot d\right) + g\left(a \cdot d\right)\chi & g\left(a \cdot d\right)\chi \\ sym & & g\left(a_3 \cdot d\right) + g\left(a \cdot d\right)\chi \end{bmatrix}. \tag{12}$$

E is Young's modulus, and ν is Poisson's ratio. Parameter $\chi = \frac{\nu}{1-2\nu}$ takes Poisson's ratio into account. For computational purposes, we introduced a and a_i as switch variables to control the effect of the damage field [16]. In tension, their value is 1, while in compression, they are set to 0.

As there is no shear strain in the principal directions, the shear term is mostly omitted from this description and can therefore only be found via numerical experimentation. We discovered that for all stable cases ($\varepsilon_1 \neq \varepsilon_2 \neq \varepsilon_3$), this block reduced to a 3-component diagonal matrix with μ being the shear modulus:

$$S_{11} = \mu\frac{|\varepsilon_1|\, g\left(a_1 \cdot d\right) + |\varepsilon_2|\, g\left(a_2 \cdot d\right)}{|\varepsilon_1| + |\varepsilon_2|}, S_{22} = \mu\frac{|\varepsilon_1|\, g\left(a_1 \cdot d\right) + |\varepsilon_3|\, g\left(a_3 \cdot d\right)}{|\varepsilon_1| + |\varepsilon_3|}, S_{33} = \mu\frac{|\varepsilon_2|\, g\left(a_2 \cdot d\right) + |\varepsilon_3|\, g\left(a_3 \cdot d\right)}{|\varepsilon_2| + |\varepsilon_3|}. \tag{13}$$

Damage affects the shear stiffness components by a linear combination of the corresponding in-plane principal strain values. As shown in Fig. 2, this solution gives indistinguishable final results for all stiffness tensor components to the original description for stable cases but also remains stable when eigenvalues are equal. Additionally, we gained 29% in computational efficiency compared to the derivative solution.

2.3. Staggered time integration

Various methods were proposed to resolve the coupled mechanical and fracture problem [20,34–37]. The present implementation solves the weakly coupled equations in a staggered manner [14,38]. Thus, when searching for the mechanical solution, the damage variable is kept constant and the energy of the displacement problem can thus be written as follows:

$$\Pi^{u} = D\left(\dot{\mathbf{u}}\right) - E\left(\mathbf{u}, d\right) + \Pi^{\text{ext}} \tag{14}$$

where Π^{ext} is the external work done by the body ($\overline{\gamma}$) and boundary ($\overline{\mathbf{t}}$) forces:

$$\Pi^{\text{ext}} = \int_{\Omega} \overline{\gamma} \cdot \mathbf{u} dV + \int_{\partial\Omega} \overline{\mathbf{t}} \cdot \mathbf{u} dA. \tag{15}$$

By taking the variation of Eq. (14), the corresponding strong form can be obtained:

$$\begin{array}{lll} \delta\Pi^u = 0 \quad \forall\delta\mathbf{u} \rightarrow \quad \nabla\boldsymbol{\sigma} + \overline{\gamma} = \rho\ddot{\mathbf{u}} & \text{in} & \Omega, \\ \boldsymbol{\sigma}\cdot\mathbf{n} = \overline{\mathbf{t}} & \text{on} & \Gamma_N, \\ \mathbf{u} = \overline{\mathbf{u}} & \text{on} & \Gamma_D, \end{array} \tag{16}$$

where $\overline{\mathbf{u}}$ is the prescribed displacement field on Γ_D. This equation system is then solved for $\boldsymbol{u}$, assuming a fixed d.

Similarly, the Lagrangian equation of the phase-field problem is obtained as:

$$\Pi^d = \int_{\Omega} \left[g_c \gamma(d, \nabla d) + g(d)\, \psi_0^+ \right] d\Omega, \tag{17}$$

where ψ_0^+ is the fixed, undamaged tensile strain energy density.

Irreversibility of the damage is ensured by the bound constrained optimization algorithm. Thus, the corresponding strong form can be expressed as:

$$\begin{array}{lll} \delta\Pi^d = 0 \quad \forall\delta d \rightarrow \quad \frac{g_c}{l_c c_\omega}\left(\frac{\partial\omega}{\partial d} - l_c^2 \Delta d\right) = \frac{\partial g}{\partial d}\psi_0^+ & \text{in} & \Omega, \\ \nabla d \cdot \mathbf{n} = 0 & \text{on} & \Gamma, \\ \dot{d} \geq 0 & \text{in} & \Omega. \end{array} \tag{18}$$

Further details about the mathematical theory of the phase-field formulation can be found in Ref. [38] and in our recent papers [14,16].

2.4. Bound constrained optimization

One of the standard approaches to enforce the irreversibility constraint ($\dot{d} \geq 0$) in AT2 models is to replace the crack-driving function (elastic tensile strain energy) in Eq. (17) with a history field [38]:

$$\mathcal{H}(\mathbf{x}, t) = \max_{s\in[0,t]} \psi_0^+(\mathbf{x}, s). \tag{19}$$

This variable corresponds to the maximum elastic energy over the full temporal history. While there is no mathematical proof that the approximation really enforces the irreversibility of the damage, it is the most widely used technique due to its simplicity. Even in our previous Abaqus implementations [14,16], we promoted Miehe's technique. However, it has to be noted that if the gradient term in Eq. (17) increases, d has to decrease for a constant $\mathcal{H}$, which violates the irreversibility condition locally.

On the other hand, this method is not applicable for AT1 models. Without elastic energy, when $\mathcal{H} = 0$, the damage field becomes negative. Therefore, Wu and Huang [20] introduced a modification to the energy history calculation:

$$\mathcal{H}(\mathbf{x}, t) = \max_{s\in[0,t]} \left\{ \hat{\psi}_0, \psi_0^+(\mathbf{x}, s) \right\}, \tag{20}$$

where $\hat{\psi}_0$ is an initial elastic energy threshold. The main advantage of this modification is that it suppresses the negative damage values and makes the history approach capable of obtaining the correct homogeneous solution. However, it overestimates the damage topology, and consequently, with a gradient present in the damage field, Wu's and Huang's method underestimates the resistance of non-homogeneous problems when the crack is defined as a Dirichlet boundary condition. Further explanation is given in Sections 3.2 and 4.1.

In this paper, we therefore enforced a positive damage increment with Lagrange multipliers. Recently, Wheeler et al. [39] proposed an augmented-Lagrangian method to solve the bound constrained optimization of the phase-field problem. This approach was later used to approximate cracks under dynamic and cohesive fracture [40]. Due to the Abaqus/UEL framework's limitations, particularly that we have no access to the global residue and stiffness arrays, the computationally slightly more expensive Lagrange multiplier technique was applied [31]. A detailed comparison of the computational cost of each method is presented in Appendix D. An interesting future project would be to test other possibilities, such as Intel MKL options or the work of Papazafeiropoulos et al. [41].

To introduce the Lagrange multipliers, the original Lagrangian equation was modified as follows:

$$\mathcal{L}^d = \Pi^d + \sum_{j=\{\mathbf{d}_n > \mathbf{d}_{n+1}\}} \lambda_j f_j(d) \tag{21}$$

where functions f_j contain the inequalities responsible for enforcing the lower bound of the nodal damage values:

$$f\left(d_{n+1}\right) = -d_{n+1} + d_n \leq 0. \tag{22}$$

In Eq. (21), the index j denotes the active constraints (where $d_n > d_{n+1}$) and λ is the Lagrange multiplier.

Fig. 3 shows the gradual convergence of the phase-field topology within one load step.

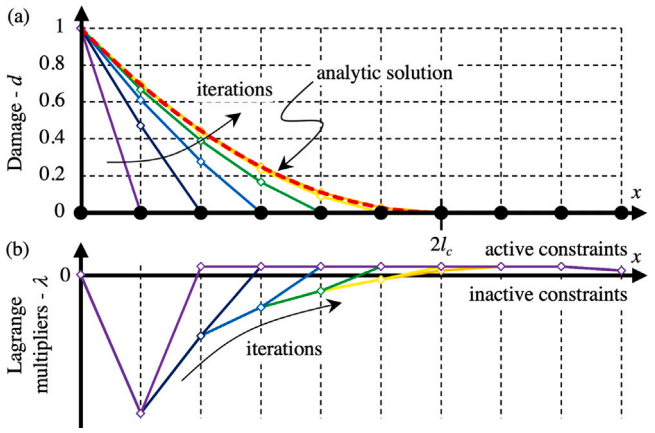

Fig. 3. (a) Crack topology in a single load step as a function of inside iterations. (b) The values of Lagrange multipliers for each iteration. When $\lambda > 0$ the irreversibility constraints are active.

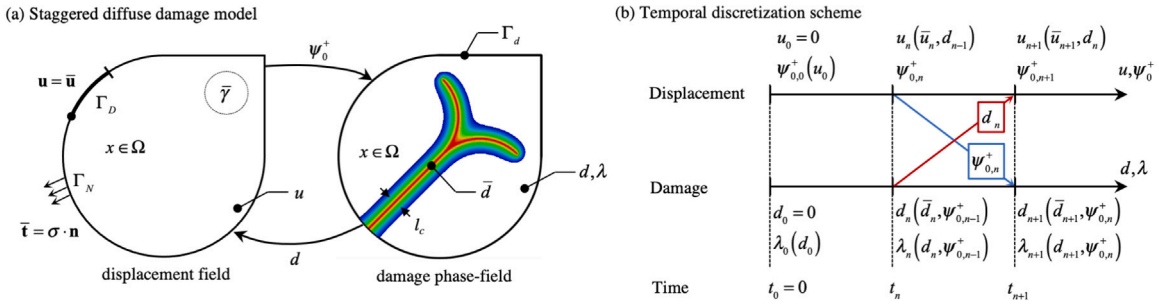

Fig. 4. Flowchart of the staggered solution used to solve the coupled mechanical and fracture problem in Abaqus.

2.5. ABAQUS implementation

Due to the limitations of the UEL option in Abaqus/Standard the two problems were solved at the same time but independently. The damage variable and the elastic strain energy were updated at the beginning of the step, and then kept constant. Fig. 4(a) shows the schematic concept of the staggered solution, while part (b) illustrates the time integration in Abaqus. The two horizontal arrows represent the two staggered problems in time. In each step, the information was exchanged between the elements only in the first one. This information contained the elastic strain energy and the damage from the previous equilibrium state. Two finite element layers were used: one for the mechanical problem with a constitutive model programmed in a UMAT, and one for the phase-field element included as a UEL subroutine. The basic concept is depicted in Fig. 5 and recommendations for developing a model in Abaqus are given in Appendix C.

The solution for both displacements, damage phase-field and Lagrange multipliers was obtained with the following Newton–Raphson iteration:

$$\begin{bmatrix} \mathbf{K}^u\left(\boldsymbol{u}_{n+1}^{(i)}, \boldsymbol{d}_n\right) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^d\left(\boldsymbol{d}_{n+1}^{(i)}, \psi_{0,n}^{+}\right) & \mathbf{A}^T \\ \mathbf{0} & \mathbf{A} & \boldsymbol{\Lambda} \end{bmatrix} \begin{bmatrix} d\mathbf{u}_{n+1}^{(i+1)} \\ d\mathbf{d}_{n+1}^{(i+1)} \\ d\lambda_{n+1}^{(i+1)} \end{bmatrix} = - \begin{bmatrix} \mathbf{r}^u\left(u_{n+1}^{(i)}\right) \\ \mathbf{r}^d\left(d_{n+1}^{(i)}, \lambda_{n+1}^{(i)}\right) \\ \mathbf{r}^\lambda\left(d_{n+1}^{(i)}, \lambda_{n+1}^{(i)}\right), \end{bmatrix} \tag{23}$$

Here, n is the time step, and i is the Newton–Raphson iteration number. The solution for each degree of freedom is gradually added up: *e.g.*, $\mathbf{d}_{n+1} = \mathbf{d}_n + \sum_i d\mathbf{d}_{n+1}^i$. For dynamic cases, the default Hilber–Hughes–Taylor (HHT) [42] time integration scheme was used. Details of the elementary matrices are presented in Appendix B.

3. Effect of numerical parameters

3.1. Homogeneous solution

The simplest way to analyze the phenomenological effect of g_c and l_c is by calculating the homogeneous solution of Eq. (18). When the gradient term is neglected, the solution becomes relatively simple, and the damage can be expressed directly as a function

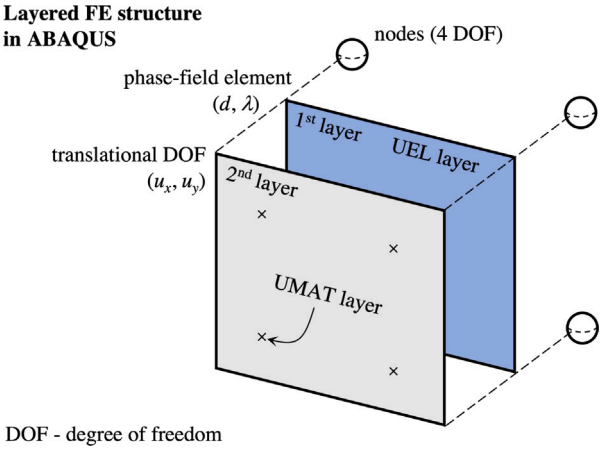

Fig. 5. 2D representation of the layered finite element structure in Abaqus. In the case of the AT1 approximation, all nodes have four degrees of freedom (DOF): two translational (u_x, u_y), damage (d), and the Lagrange multiplier (λ). The first (UEL) element contributes to the damage problem with the additional of the Lagrange multipliers enforcing the inequality conditions. The second (UMAT) element, on the other hand, solves the mechanical problem.

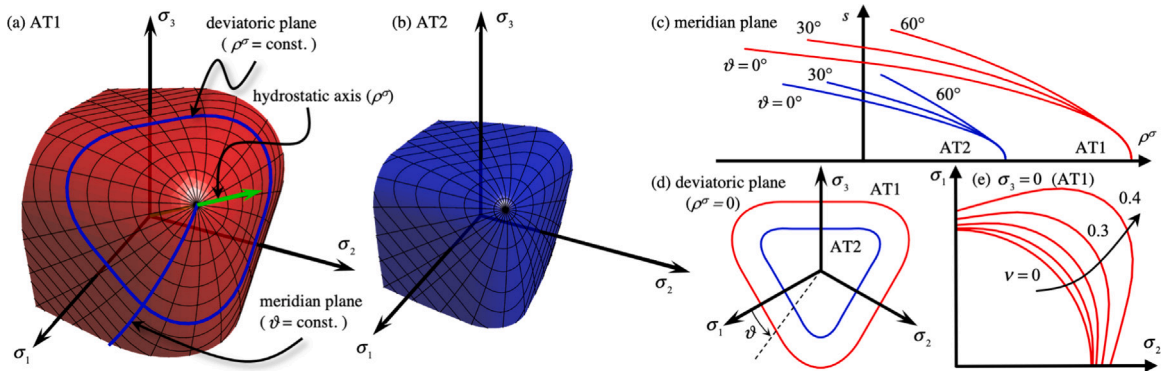

Fig. 6. Homogeneous solution for problem AT1 and AT2: (a) and (b) failure surfaces in the space of principal stresses; (c) meridian slices for different Lode angles; (d) deviatoric slices at zero hydrostatic stress; (e) AT1 model for in-plane stress conditions and different values of Poisson's ratio.

of the elastic strain energy. The most frequently shown result is the stress–strain curve under uniaxial tension [14,43]. However, as demonstrated earlier [44], the maximum stress achievable during deformation is affected by the ratio of the principal stresses.

Recent atomic-scale simulations exhibited [45] that the response of a brittle material (*e.g.,* silica) subjected to tension might depend not only on the hydrostatic stress state but also on the Lode angle. Therefore, the resistance of a material should be described with a three-dimensional failure/yield surface.

Fig. 6(a,b) display the surface of failure in the space of all three principal stresses for AT1 and AT2 models for the same phase-field parameters. Fig. 6(c–e) depict slices of the meridian planes (constant Lode angle), deviatoric planes (constant hydrostatic stress), and in plane stress ($\sigma_3 = 0$) conditions. The plotted surfaces resembled a Rankine criterion [46] with rounded corners. Indeed, when $\nu = 0$, the phase-field solution recovered a Rankine-type surface with circularly rounded edges in the tensile domains.

Since the AT2 model had no elastic threshold, it was consistently inferior in strength to the AT1 formulation. Moreover, the tensile meridians were weaker than their compression counterparts.

The algorithm calculating the maximum stress as a function of Poisson's ratio and the principal stress state is available as a Supplementary File. The homogeneous solution can be used to give an initial approximation for l_c as a function of the measured material strength [44].

3.2. Unidirectional damage topology

To demonstrate the validity of the bound constrained optimization scheme, Fig. 7 displays the AT1 phase-field solution of a 1D boundary problem. The regularization length was fixed at $l_c = 3$ mm and a Dirichlet condition of $\hat{d} = 1$ was enforced at $x = 0$, while the elements remained undeformed ($\psi_0^+(x) = 0$). Since the crack-driving force became negative when no elastic energy was applied for AT1 models, negative damage values could appear without the specialized treatment. This phenomenon is a violation of phase-field principles and although Wu's [20] idea (with the initial crack driving force shown in Eq. (20)) solves this problem, it

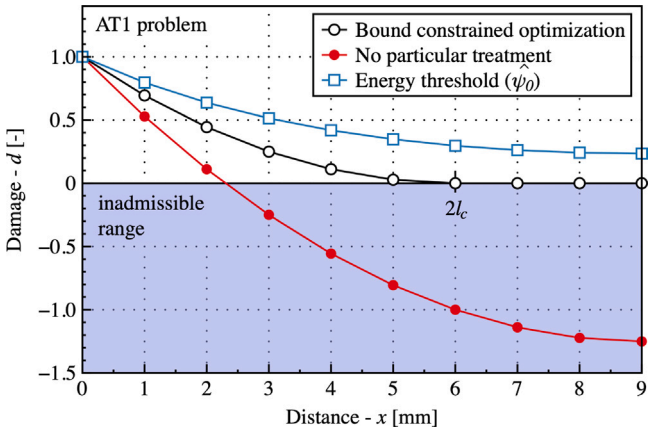

Fig. 7. Equilibrium phase-field topology with $\hat{d} = 1$ at $x = 0$ mm. Each symbol represent a nodal value of the finite element calculation.

introduces a $d \neq 0$ condition everywhere in the model. When the phase-field is initiated only by mechanical deformation (without the Dirichlet condition), this phenomenon is not present.

Therefore, the second description generates more fracture energy than the theoretical value, thereby significantly underestimating the resistance of the sample significantly (see later in Section 4.1).

The bound constrained optimization scheme adapted in this paper, however, converges to the analytical solution. Ergo, values of the damage field vary according to a second-order polynomial function until $x = 2l_c$, and then maintain the value of zero for the rest of the model.

To avoid instability and to satisfy the Ladyzhenskaya–Babuška–Brezzi condition, we paid particular attention to the fact that there were fewer Lagrange multipliers active in the model than other degrees of freedom [47,48]. As the energy field is fixed, we have always experienced a monotonous convergence in the multiplier space to the final solution. Hence, we did not encounter instabilities.

3.3. Effective fracture toughness

The finite element approximation of the phase-field solution used to model problems in fracture mechanics suffers from two major types of errors. The loading history has to be divided into sub-steps [49], and the simulation space has to be meshed with finite-size elements. Therefore, both the temporal and spatial discretization affect the final results. In our previous papers [16,44], we tackled the issue of the time step control to minimize the error introduced by the staggered algorithm. Here, we elaborate on the second type of error introduced by the spatial tessellation.

The inaccuracies originating from the spatial discretization can further be divided into two categories: (i) approximation and (ii) localization errors. To demonstrate these cases, we prepared a simple unidirectional test with finite elements in a single row. In the first step, a phase-field value of $\hat{d} = 1$ was defined at $x = 0$. Then, the sample was extended in the x direction to analyze the evolution of the damage topology. The analytical solution of the theoretical crack surface is well-known. Depending on the symmetry of the localization, it can be approximated with finite elements by calculating W in Eq. (4). This section compares the analytical and the approximated crack surfaces to obtain the error.

Fig. 8(a) depicts the difference between the analytical solution and a crude finite element approximation for an AT1 topological solution. No deformation was applied and a finite element mesh size of l_c was used to magnify the difference shown in light blue. Fig. 8(b) and (c) show the increment in crack surface introduced by the localization for symmetric boundary conditions and for a complete model, respectively.

The phase-field values were calculated as nodal degrees of freedom, but were taken into account phenomenologically at the integration points. Consequently, with the topological solution shown in part (a), the crack could not open freely as the element itself still retained most of its stiffness and could carry a load. As a consequence, when further deformation was applied, the phase-field created a region where the crack could localize. This only took place when the phase-field was constant in the element and all of its nodal damage values were close to 1. However, at this stage the crack surface was increased by a well-defined, mesh-dependent amount.

It is clearly visible from Fig. 9 that the localization dominated the error introduced by the finite element approximation. To quantify the ratio between the two problems, Fig. 9 shows the relative difference between the analytical crack surface and the finite element approximation.

Even if the approximation error increased with a second-order polynomial, up to $l_c/h = 1$ the localization error is several orders of magnitude higher than the approximation error.

The existence of the approximation [34] and the localization [50–52] errors is well-known in the literature. To the best of our knowledge, a graphical explanation for the solution is still missing. For the most part, two solutions are employed: (i) to reduce the

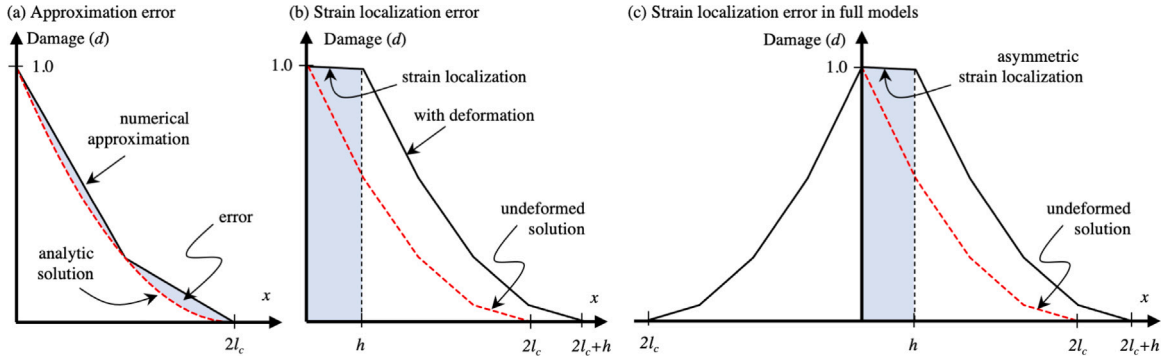

Fig. 8. Damage profiles for different cases. Part (a) highlights the difference in blue between the numerical approximation and the analytic solution for the AT1 model. In part (a), for demonstration purposes, we show the case with $l_c/h = 1$. In parts (b) and (c), the increase in damage is depicted due to strain localization with a symmetry boundary condition and for full models, respectively. The size of the finite elements is shown by h, and $l_c/h = 2$.

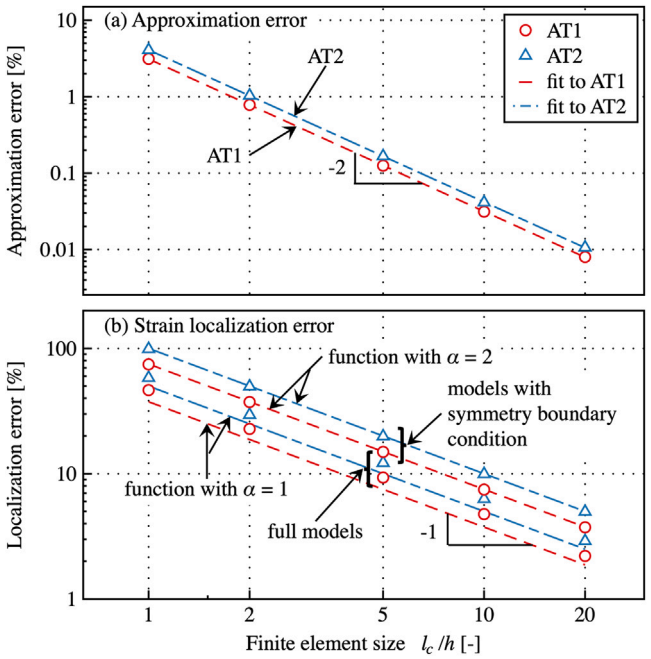

Fig. 9. Error introduced by the finite element implementation for both AT1 and AT2 models. Part (a) shows the error introduced by the discrete approximation. The numerical data was fitted with the function $\zeta_0\left(\frac{l_c}{h}\right)^{-2}$, where $\zeta_0 = 3.1129$ for AT1, and $\zeta_0 = 4.0892$ for the AT2 models. Part (b) depicts the error generated by the strain localization in a single element. The numerical results are fitted using function (25) with $\alpha = 1$ for the full model and $\alpha = 2$ for the models with symmetry boundary conditions.

effect of the approximation error, the finite element size should be smaller than $l_c/2$ [34]; (ii) to account for the localization an effective (numeric) fracture toughness (g_c) is introduced [50,52]:

$$g_c = g_{mat}\left(1 + \alpha\frac{h}{c_\omega l_c}\right)^{-1} = \frac{g_{mat}}{1+\zeta}. \tag{24}$$

Here, ζ accounts for the surplus fracture energy:

$$\zeta = \alpha\frac{h}{c_\omega l_c}, \tag{25}$$

and g_{mat} is the material's real fracture toughness. Furthermore, $\alpha = 1$ for a full model and 2 for a geometry with symmetry boundary conditions, and h is the characteristic size of the finite element perpendicular to the crack propagation. This approximation results in an equivalent crack surface ($\hat{\Gamma}$) through the following normalization:

$$\hat{\Gamma} = \frac{\Gamma}{1+\zeta} \tag{26}$$

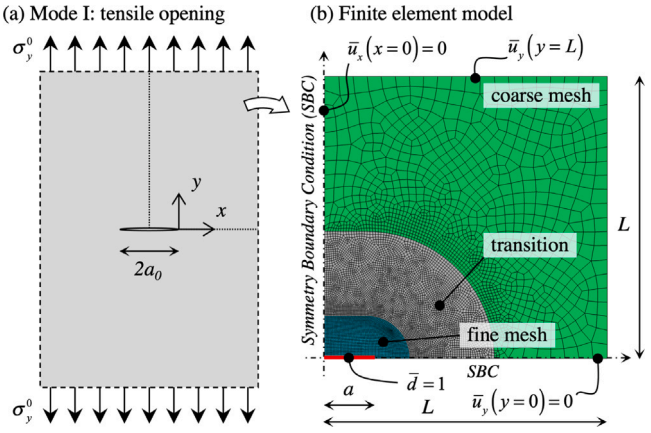

Fig. 10. Mode I tensile opening: (a) schematic problem; (b) finite element model with double symmetry: $a_0 = 2$ mm and $L = 100$ mm.

4. Benchmark tests and numerical examples

In order to verify the newly implemented model, this section presents well-known benchmark tests with increasing complexity. First, Section 4.1 illustrates through a simple extension (mode I) test how the critical loading converges to the theoretical solution of Griffith [8]. Next, Section 4.2 shows the capabilities of the damage-gradient approach to model mixed-mode fracture in 3D (mode I + III). Our third example in Section 4.3 is meant to display that the modified energy decomposition produces symmetric results in dynamic instabilities despite the randomized mesh. Finally, in Section 4.4, the last case demonstrates how close a phase-field simulation can reproduce experimental crack patterns in a Kalthoff–Winkler test.

4.1. Mode I tensile opening

This elementary case was recently studied in connection with the coupled criterion [44], where a theoretical explanation was provided for the regularization parameter l_c with a connection to the material's tensile strength. In that paper, we demonstrated that if l_c tends to zero in an AT2 model, the critical loading slightly overestimates (≈10%) the model's resistance compared with Griffith's theoretical solution [8]:

$$\sigma_y^{0,G} = \sqrt{\frac{8 g_{mat} \mu}{(\kappa + 1)\, a_0 \pi}}, \tag{27}$$

where μ is the shear modulus and $\kappa = 3 - 4\nu$ for the plane strain case.

On the other hand, Kristensen et al. [43] demonstrated that AT1 models with the energy threshold criterion had a tendency to underestimate this value. To tackle this elementary conundrum, a finite element model (FEM) depicted in Fig. 10 was developed. Only one quarter of the space was modeled with symmetric boundary conditions on the middle lines. The sample's overall length (L) was taken as $50a_0$ to minimize the effect of the finite size. Dirichlet boundary conditions were prescribed on the edges as displacements and a $d = 1$ phase-field value on the crack surface. Young's modulus was set to 3 GPa with a Poisson's ratio of 0.37 and a critical energy release rate of $g_c = 300$ J/m^2. The regularization parameter and the finite element mesh were varied throughout the study.

The automatic time integration scheme was applied to identify the maximum reaction force with an energy constraint of $\eta = 50$ % [16,44]. The maximum reaction force on the top edge was then divided by the length of the model (L) to obtain $\sigma_y^{0,c}$.

Fig. 11 presents the critical loading values as a function of l_c/h with and without correction (Eq. (24)) for both the AT1 and AT2 models. It is visible that without the energy correction, the ratio l_c/h should be at least 5–6. This is in good correspondence to other phase-field simulations [22,43]. The results demonstrate that the effective fracture toughness (see Eq. (24)) showed great potential for improving the localization error. Even with a relatively coarse mesh ($l_c/h = 1$), the difference was less than 3% (for AT1) and 1% (for AT2) compared to the convergence limit (without correction). The remaining difference can be accounted for by the approximation error discussed in Section 3.3.

Eq. (27) suggests that the critical load should depend on the toughness via a square root function. In the inset of Fig. 12, we plotted $\sigma_y^{0,c}$ calculated by phase-field simulations as a function of g_c. It can be seen, that a square root function fit the obtained results perfectly. With this in mind, the critical loading could be plotted as a function of l_c. Fig. 12 shows that both the thermodynamically consistent AT1 (present implementation) and AT2 models converged to the theoretical solution, while the AT1 solution with the energy threshold underestimated the resistance of the model ≈10%.

Assigning a physical meaning to l_c is still an ongoing debate. Recently, we used a long-known size effect [53] to study the impact of l_c on the tensile strength of the material. We observed that when a large initial crack was present, the reduction of l_c let

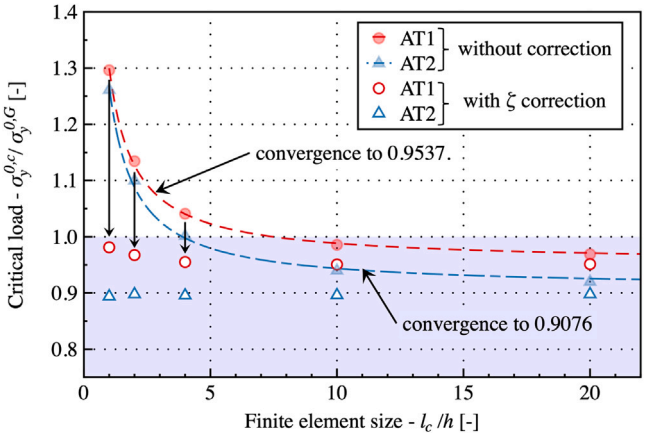

Fig. 11. Critical stress as a function of the finite element size. The length scale parameter was kept constant at $l_c = 0.2$ mm. The stress values were normalized by Griffith's solution (Eq. (27)). Empty symbols show the results with ζ correction (Eq. (24)), whereas, in case of the transparent filled ones, $g_c = g_{mat}$. The curves present a hyperbolic fit on the uncorrected results, with convergence to respectively 0.9537 and 0.9076 for AT1 and AT2 when $l_c/h \to \infty$.

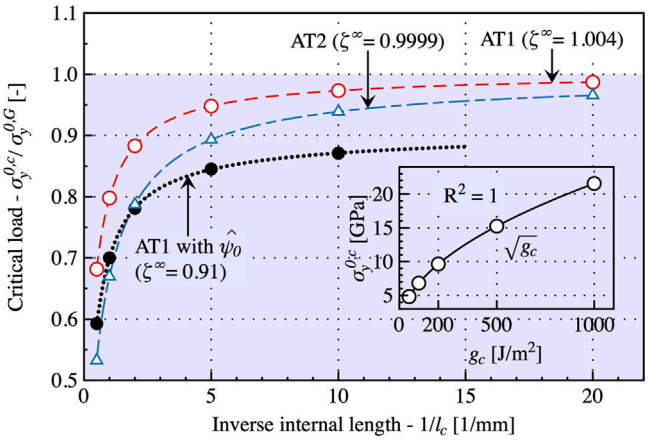

Fig. 12. Normalized critical loading as a function of $1/l_c$ for different models. At $l_c \to 0$ the functions tended to ζ^∞.

us converge to the finite value proposed by Griffith. If the defect size was reduced to zero, the homogeneous solution, presented in Section 3.1 resulted in an infinite strength value. Fig. 13 shows this particular transition as a function of a_0. It can be seen that both models corresponded well to their homogeneous solutions for very small a_0. Conversely, they converged to Griffith's theory (dashed line) when a_0 was much larger than l_c.

4.2. Mixed mode I and III

Cracks subjected to extension and perpendicular shear (mode I+III) initiate and propagate in a complex multi-scale manner [54]. At the small scale, near the crack tip, the mother crack fragments into numerous daughter cracks (known as *echelon* cracks) [55–57]. At the large scale, on the other hand, these daughter cracks coalesce [58,59] and the primary crack front starts to rotate around its original axis [55]. This problem was studied comprehensively in both experiments [55–57,60,61], theoretical calculations [62,63], and numerical simulations [55,59,64]. While the coalescence and the macroscopic rotation seemed easy to reproduce with the phase-field approach, the microscopic initiation was hard to induce without a fairly inhomogeneous model [64].

In this benchmark example, we tackled the two scales with two different models. To study the effect of a crack at an angle, the model depicted in Fig. 14(a) was developed. The size of the sample was taken to be $L = 40$ mm, $T = 20$ mm and $H = 60$ mm. The crack was initially placed at a 45° angle with an initial length of $a_0 = 5$ mm and the loading was applied in the x direction. The random finite element mesh of 1 250 000 tetrahedrons was generated with an average size of 0.5 mm.

To study the initiation of the echelon cracks, the model shown in Fig. 14(b) was created. The radius of the sample was $R = 0.5$ mm, while the length was $l = 2$ mm. The random mesh of 130 000 linear brick elements was generated with an average size of 0.025 mm.

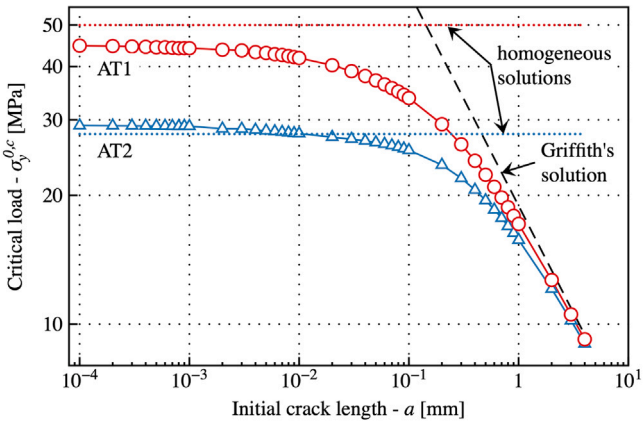

Fig. 13. Critical loading as a function of the initial crack length for the AT1 and AT2 models. The regularization length scale (l_c) was set to 0.2 mm.

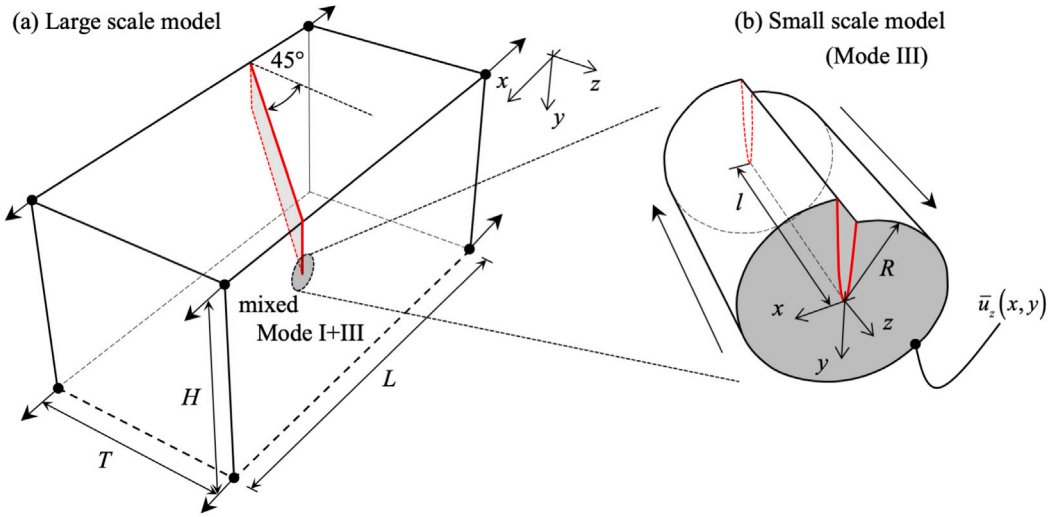

Fig. 14. Mixed mode crack propagation in extension and torsion. The problem was modeled at two scales: (a) at a large scale with mixed mode fracture; (b) at the crack tip to test the elementary initiation process.

The analytical displacements ($\bar{u}_z$) were prescribed on the mantle and on both sides of the cylinder [65]:

$$\bar{u}_z(x,y) = \frac{\tau_{yz}^0}{\mu}\sqrt{r_1 r_2}\sin\left(\frac{\theta_1+\theta_2}{2}\right), \tag{28}$$

where $r_1 = \sqrt{x^2+y^2}$, $r_2 = \sqrt{(x+2a_0)^2+y^2}$, $\theta_1 = \tan^{-1}\left(\frac{y}{x}\right)$ and $\theta_2 = \tan^{-1}\left(\frac{y}{x+2a_0}\right)$, while the damage field was kept periodic in the z direction at all times.

In all models the initial crack was represented geometrically and with a predefined Dirichlet boundary condition in the phase-field. For both models, a Young's modulus of 3 GPa, a Poisson's ratio of 0.37 and a g_c of 300 J/m^2 were used. The regularization length (l_c) was varied: : a value of 1 mm was set in the large scale model, and a value of 0.025 mm in the small scale model. The time step was controlled automatically.

Fig. 15 shows the results obtained at the large scale at various loading stages. The crack at a 45° angle was initially loaded in both tension (mode I) and out-of-plane shear (mode III). Therefore, the crack gradually twisted around the y axis to reach the perfect tensile mode, as it was energetically more favorable. As shown by experiments [55], the crack always initiated at the sides with pointy crack faces. This was clearly visible in our phase-field simulation as well, as can be seen in the first and second frames of Fig. 15. At this scale, due to the regularization, the small daughter cracks were not visible. However, according to Fig. 16, the localized crack planes for a pure Mode III deformation showed the elementary initiation process.

As reported by Pham and Ravi-Chandar [64], the deviation from the volumetric damage propagation in phase-field in Mode III is complicated. Particularly when there are no initial imperfections, the phase-field simulation produces a concentric (volumetric) damage field. To find the inclined planar cracks (as seen in experiments), Ref. [64] introduced a large number of small, randomly

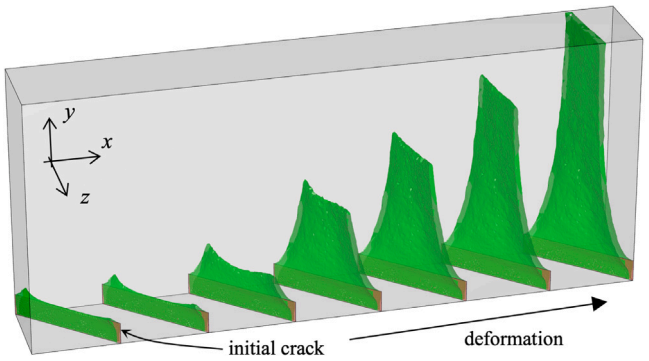

Fig. 15. Crack pattern for the large scale model (isosurface at $d = 0.8$) as a function of deformation.

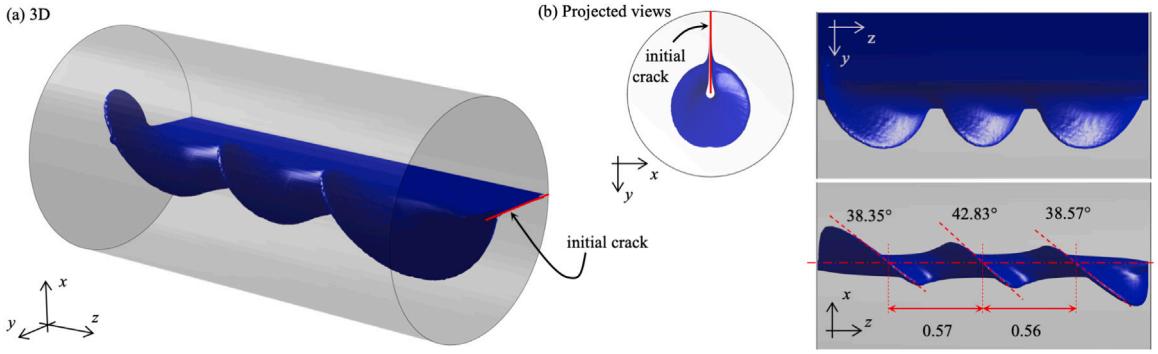

Fig. 16. Crack pattern for the small scale model (isosurface at $d = 0.8$) at different views.

distributed damaged zones throughout the sample. They showed that these initial imperfections had a significant effect on the crack patterns. We were able to confirm the insufficiency of only a few initial defects. Therefore, this paper emphasizes the importance of the constraint damage evolution (*e.g.,* inhomogeneous toughness, unequal boundary conditions) and we hence we created a model where the prescribed displacements were enforced not only on the mantle of the cylinder, but also on the sides. As the analytical (singular) displacement field deviated slightly close to the regularized one, the damage evolution was disturbed near the sides at the crack tip. This gave us just enough perturbation in the phase-field to induce the localization as seen in experiments and such an initial localization was enough for the helical daughter cracks to initiate and propagate gradually in a periodic manner.

However, we emphasize that the present model was only proof of concept to demonstrate the feasibility of this inclined localization in phase-field simulations. It would be crucial to come up with a physically sound idea to study these instabilities, *e.g.,* with the randomly varying fracture toughness. However, this study exceeds the scope of the present article.

4.3. Dynamic branching

The geometry of the dynamic tensile sample is depicted in Fig. 17(a). The finite element mesh was created with a randomly generated 313 000 elements with a maximum size of 0.125 mm. Young's modulus was set to 32 GPa with a Poisson's ratio of 0.2 and a density of 2 450 kg/m^2. The fracture parameters were $g_c = 3$ J/m^2 and $l_c = 0.5$ mm. A fixed time step size of $\Delta t = 10^{-8}$ s was defined with the default $\alpha_{\text{HHT}} = -0.05$ damping parameter. The applied tension on the top and bottom sides were uniformly $\sigma_y = 1$ MPa.

The two crack patterns in Fig. 17(b) indicate that the branches appeared symmetrically with both AT1 and AT2 models. It can be seen that the AT1 model confined the damage in a smaller region, however to dissipate the same amount of energy, the peak where d ≈ 1 was much wider. This result was a correction to the previous implementation, where a small perturbation was used to stabilize the energy decomposition scheme. Interestingly, when a structured mesh was applied, the branches remained symmetric. Therefore, we attributed this instability to the irregularity of the mesh. However, the present results revealed that the energy decomposition with a small perturbation could cause artificial results in sensitive cases. It is therefore crucial to use a numerically stable energy decomposition scheme. Furthermore, the added damage diffusion is intrinsically related to the localization error and mesh anisotropy effects [16]. However, the detailed analysis would be essential but it exceeds the aims of present article.

For quantitative comparison, the crack tip velocity is shown in Fig. 17(c) for both models. The position of the crack tip was calculated as the largest extension of the phase-field where its value was larger than 0.9. The maximum velocity is in good corresponded well with the theoretical limit [66].

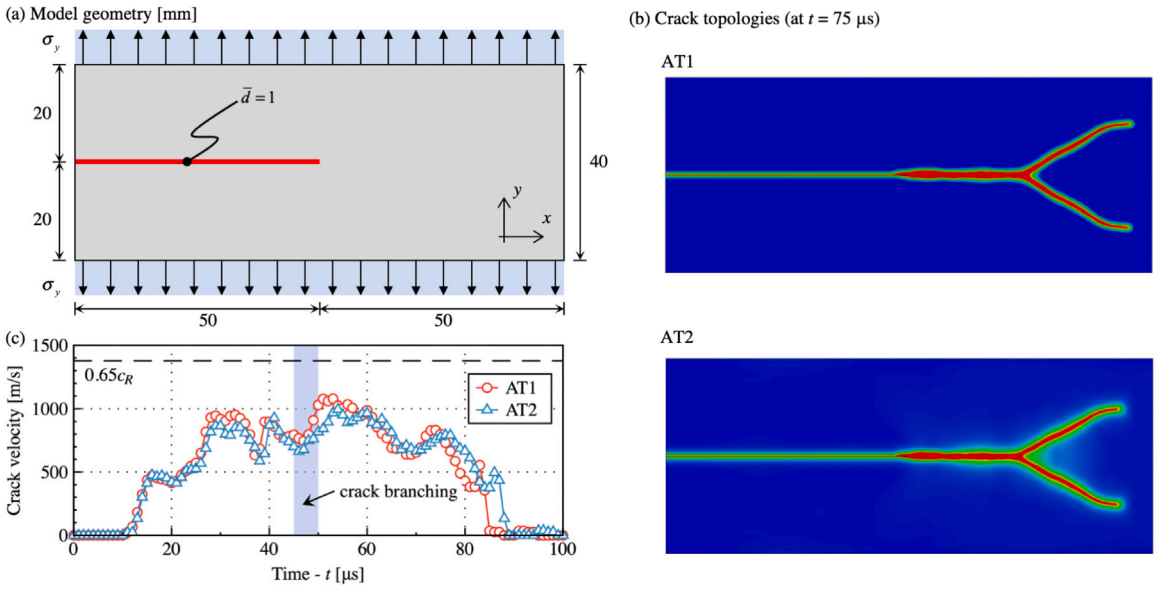

Fig. 17. Rectangular plate subjected to uniform unidirectional dynamic traction: (a) Geometry with dimensions in mm; (b) Crack pattern at 75 μs for AT1 and AT2 models; (c) Crack tip velocity as a function of time.

4.4. Kalthoff–Winkler test

We finally investigated the test proposed by Kalthoff and Winkler [67], but with slight modifications to the geometry and the material to avoid shear banding. The geometry of the specimen is depicted in Fig. 18(a). In the phase-field simulation, only one-half of the plate was modeled with symmetric boundary conditions at the middle line. An impact velocity of 25 m/s was chosen, and in the FEM, the velocity imposed on the impact surface was constant after an initial linear increase of 10^{-6} s. The material properties $E = 6$ GPa, $\nu = 0.35$, $\rho = 1180$ kg/m^3, $l_c = 0.4$ mm, and $g_c = 600$ J/m^2 were independent of the deformation rate and the mesh was refined with a maximum size of 0.2 mm in the area where the crack was expected to evolve. This way, a model of 83 000 finite elements was generated. The time step was fixed at 10^{-8} s.

Fig. 18(b) depicts the crack (with $d > 0.95$) in the undeformed configuration for the AT1 and AT2 models. Fig. 18(c) shows the experimental crack pattern obtained by Rian Seghir and Julien Réthoré at École Centrale de Nantes (unpublished results). For the test, 10 mm thick PMMA samples were used with an impactor arriving at a speed of 25 m/s. This paper focused only on the *ex situ* crack pattern in comparison to the simulations, so the detailed analysis of the test will be part of a separate paper.

There were three distinct characteristics of the post-experimental crack pattern, which appeared consistently: (a) the initiation angle; (b) the position of the branching; and (c) the secondary crack initiating from the tensile (bottom-left) side. It can be seen from the numerical comparison that the initiation angle in the phase-field simulations was comparable to the experimental measurement. Only a few degrees of difference can be observed. The branching appeared slightly later in the simulations, which might indicate a somewhat lower g_{mat} in reality. High-speed recordings confirmed that the secondary crack indeed initiated from the left side and propagated in the direction of the main crack, as seen in the simulations. In conclusion, all three observations were in agreement with the experimental tests.

The secondary branching observed in the AT2 simulation was an instability, which required studying, as there were some experimental samples where the main crack branched a second time on the left, but this phenomenon was not reproduced in every case.

5. Discussion

One of the most significant advantages of the phase-field formulation for brittle fracture is that the method does not need any primary crack definition. However, while localization can appear automatically, without an initial Dirichlet boundary condition, the critical load is overestimated compared with the analytical (singular) solution. Without an initially damaged zone, an energy cost is associated with the initiation process. Simply put, to create an equivalent crack length in an undamaged solid, we need energy, and therefore more loading can be applied. This phenomenon was shown to overestimate the resistance even if a bound constraint optimization is used to enforce positive damage evolution by more than 20% [43,52,68].

On the other hand, the energy threshold solution underestimates the critical loading due to the larger crack surface. Interestingly, these two effects were shown to counteract each other and give results relatively close to the analytical formulation [43]. Therefore,

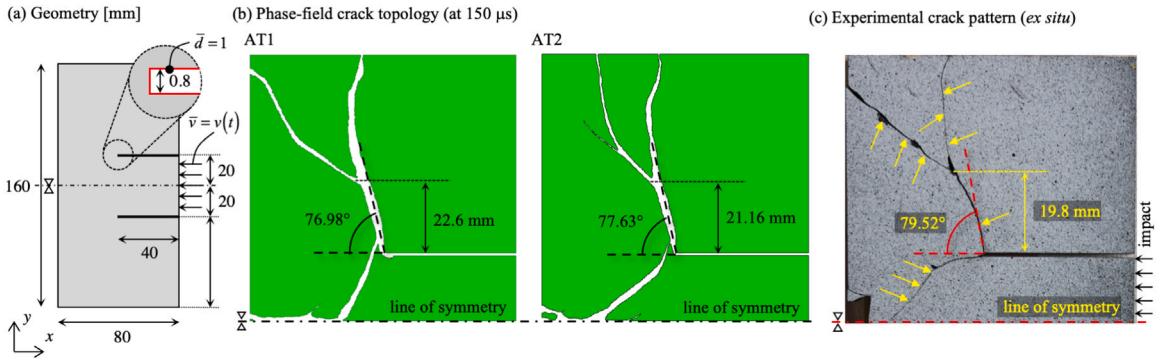

Fig. 18. Kalthoff–Winkler test: (a) geometry with dimensions in mm; (b) crack pattern after 150 μs (elements with $d > 0.95$ are not shown); (c) *ex situ* experimental crack pattern on PMMA.

by omitting the Dirichlet boundaries in the phase-field, a computationally simplified solution would be well equipped in some practical cases.

The present research showed that initial defects and a small perturbation in the phase-field could also experimentally reproduce the observed results. This gives rise to the issue of how to introduce instabilities in the crack front to deviate from the volumetric phase-field solution and localize the crack front in a plane, as seen in the mixed mode I+III tests. Maybe introducing a spatially varying toughness field or an initial damage distribution [64] would solve the problem.

Recently, Kumar [69] proposed the addition of a phenomenological crack driving force, which can replace l_c in representing the strength of the material. This method would be promising to solve both nucleation and localization problems. The only issue is that it introduces another set of material-dependent parameters, which need to be calibrated experimentally. Therefore, it is crucial to experimentally study damage nucleation, crack initiation, and propagation and to compare local results to phase-field calculations.

This leads us to the essential question, whether to use AT1 or AT2 models (or other types of mixed implementations [30]). Initially, AT2 models were proposed due to their simplicity. However, since more advanced numerical formulations have appeared, it seems that AT2 models are inferior to their counterparts with elastic thresholds. We argue that this is only a matter of scale and material: it was for instance shown in microscopic samples [70] that even for materials as brittle as silica, it was hard to define a pure, elastic regime.

6. Conclusion

This paper presents the first thermodynamically consistent Abaqus implementation of the linear gradient damage model. To this aim, we used a bound constraint optimization scheme with Lagrange multipliers to enforce positive damage fields and damage irreversibility. The code was implemented in a combined UEL and UMAT structure and could solve both 2D and 3D problems, as well as static and dynamic ones. We updated the spectral decomposition scheme of Bernard et al. [33] to remain stable even in originally unstable cases. Furthermore, we presented a graphical explanation of the spatial errors manifesting in the phase-field fracture approximation.

We observed that both the AT1 and AT2 models were consistent with the analytical predictions of Griffith. In particular, when l_c approached 0, the critical loading converged to the singular solution. The stabilization of the energy decomposition scheme allowed us to obtain symmetric results in dynamic instabilities even with random finite element meshes. Finally, we found good correspondence between simulation results and experiments in both static and dynamic examples even if some examples can exhibit second order differences (like the Kalthoff–Winkler test).

AT1 models provide an initial elastic threshold and for this reason the engineering community is very eager to see their implementation. However, special treatment is required to treat negative damage values. The Lagrange multiplier technique makes it possible to solve complex optimization problems in Abaqus by adding a new degree of freedom. Combining the bound constraint optimization scheme and the AT1 representation makes it possible for the end-user to efficiently solve complex fracture problems.

Despite their usefulness, the results presented in this paper raise multiple questions, which need to be answered in the future, such as crack initiation in mixed mode deformation, the rate-dependent nature of the dynamic fracture, or the microscopic meaning of l_c.

CRediT authorship contribution statement

Gergely Molnár: Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Validation. **Aurélien Doitrand:** Conceptualization, Formal analysis, Investigation, Writing – review & editing. **Adrien Jaccon:** Investigation. **Benoit Prabel:** Funding acquisition, Investigation, Supervision, Writing – review & editing. **Anthony Gravouil:** Writing – review & editing, Supervision, Investigation.

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

The authors thank Julien Réthoré and Rian Seghir for sharing the results of the Kalthoff–Winkler experiment.

Appendix A. Rotation matrices

The strain tensor ($\boldsymbol{\varepsilon}$) is a symmetric second-order tensor and can always be represented by a 3 by 3 matrix. This tensor has three eigenvalues (ε_i) and 3 mutually perpendicular eigenvectors ($\mathbf{v}_i$). The principal strain components can be placed in the diagonal matrix ($\hat{\boldsymbol{\varepsilon}}$) so that

$$\boldsymbol{\varepsilon} = \mathbf{Q}\hat{\boldsymbol{\varepsilon}}\mathbf{Q}^T, \tag{A.1}$$

where $\mathbf{Q}$ is a rotation matrix with the eigenvectors in the same order as the eigenvalues: $\mathbf{Q} = \left[\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_3\right]$.

However, in the implemented algorithm we use Voigt's notation, where the stress and strain tensors are represented by 6 component vectors. Consequently, to execute the rotation of the stiffness matrix, the following two rotation matrices are defined:

$$\mathbf{T}_\varepsilon = \begin{bmatrix} Q_{11}^2 & Q_{21}^2 & Q_{31}^2 & Q_{21}Q_{11} & Q_{31}Q_{11} & Q_{21}Q_{31} \\ Q_{12}^2 & Q_{22}^2 & Q_{32}^2 & Q_{22}Q_{12} & Q_{32}Q_{12} & Q_{22}Q_{32} \\ Q_{13}^2 & Q_{23}^2 & Q_{33}^2 & Q_{23}Q_{13} & Q_{33}Q_{13} & Q_{23}Q_{33} \\ 2Q_{12}Q_{11} & 2Q_{22}Q_{21} & 2Q_{32}Q_{31} & Q_{22}Q_{11}+Q_{12}Q_{21} & Q_{31}Q_{12}+Q_{11}Q_{32} & Q_{22}Q_{31}+Q_{32}Q_{21} \\ 2Q_{13}Q_{11} & 2Q_{23}Q_{21} & 2Q_{33}Q_{31} & Q_{11}Q_{23}+Q_{21}Q_{13} & Q_{31}Q_{13}+Q_{11}Q_{33} & Q_{23}Q_{31}+Q_{33}Q_{21} \\ 2Q_{13}Q_{12} & 2Q_{23}Q_{22} & 2Q_{33}Q_{32} & Q_{22}Q_{13}+Q_{23}Q_{12} & Q_{13}Q_{32}+Q_{12}Q_{33} & Q_{23}Q_{32}+Q_{33}Q_{22}. \end{bmatrix} \tag{A.2}$$

and

$$\mathbf{T}_\sigma^{-1} = \begin{bmatrix} Q_{11}^2 & Q_{12}^2 & Q_{13}^2 & 2Q_{12}Q_{11} & 2Q_{13}Q_{11} & 2Q_{13}Q_{12} \\ Q_{21}^2 & Q_{22}^2 & Q_{23}^2 & 2Q_{22}Q_{21} & 2Q_{23}Q_{21} & 2Q_{23}Q_{22} \\ Q_{31}^2 & Q_{32}^2 & Q_{33}^2 & 2Q_{32}Q_{31} & 2Q_{33}Q_{31} & 2Q_{33}Q_{32} \\ Q_{21}Q_{11} & Q_{22}Q_{12} & Q_{23}Q_{13} & Q_{22}Q_{11}+Q_{12}Q_{21} & Q_{21}Q_{13}+Q_{11}Q_{23} & Q_{22}Q_{13}+Q_{23}Q_{12} \\ Q_{31}Q_{11} & Q_{32}Q_{12} & Q_{33}Q_{13} & Q_{11}Q_{32}+Q_{12}Q_{31} & Q_{31}Q_{13}+Q_{11}Q_{33} & Q_{32}Q_{13}+Q_{33}Q_{12} \\ Q_{21}Q_{31} & Q_{22}Q_{32} & Q_{23}Q_{33} & Q_{22}Q_{31}+Q_{21}Q_{32} & Q_{31}Q_{23}+Q_{21}Q_{33} & Q_{23}Q_{32}+Q_{33}Q_{22}. \end{bmatrix} \tag{A.3}$$

It is important to note that $\mathbf{T}_\varepsilon \neq \mathbf{T}_\sigma$ due to the engineering shear strain components.

Appendix B. Elementary matrices and residues

To solve the phase-field fracture problem, a split scheme operator is applied. Equation system (23) is solved iteratively using the Newton–Raphson method, where both the residue and the stiffness matrix are recalculated in each iteration.

The classic matrices are given in previous papers [14,16]:

$$r_j^u = \frac{\partial \mathcal{L}}{\partial u_j}, K_{jk}^u = \frac{\partial \mathcal{L}}{\partial u_j \partial u_k}, K_{jk}^d = \frac{\partial \mathcal{L}}{\partial d_j \partial d_k}, \tag{B.1}$$

where r_j^u is the residue of the mechanical problem for the jth degree of freedom. Similarly, K_{jk}^u and K_{jk}^d belong to the jth and kth degree of freedom of the mechanical and phase-field problems, respectively.

This paper details only the modifications necessary to enforce irreversibility with the Lagrangian multipliers. The modified residue vectors are calculated by differentiation of the modified Lagrangian function in Eq. (21):

$$\begin{aligned} r_j^d &= \frac{\partial \mathcal{L}}{\partial d_j} = \begin{cases} \frac{\partial \Pi^d\left(d_{n+1}^{(i)}\right)}{\partial d_j} - \lambda_{j,n+1}^{(i)} & \text{if} \quad d_{j,n+1}^{(i)} \le d_{j,n} \quad \text{or} \quad \lambda_{j,n+1}^{(i)} \ge 0, \\ \frac{\partial \Pi^d\left(d_{n+1}^{(i)}\right)}{\partial d_j} & \text{otherwise}, \end{cases} \\ r_j^\lambda &= \frac{\partial \mathcal{L}}{\partial \lambda_j} = \begin{cases} -d_{j,n+1}^{(i)} + d_{j,n} & \text{if} \quad d_{j,n+1}^{(i)} \le d_{j,n} \quad \text{or} \quad \lambda_{j,n+1}^{(i)} \ge 0, \\ 0 & \text{otherwise}. \end{cases} \end{aligned} \tag{B.2}$$

While the elementary stiffness components are obtained by taking the second derivatives:

$$
\begin{aligned}
A_{jj} &= \frac{\partial^2 \mathcal{L}}{\partial d_j \partial \lambda_j} = \begin{cases} -1 & \text{if} \quad d^{(i)}_{j,n+1} \le d_{j,n} \quad \text{or} \quad \lambda^{(i)}_{j,n+1} \ge 0, \\ 0 & \text{otherwise,} \end{cases} \\
A_{jk} &= \frac{\partial^2 \mathcal{L}}{\partial d_j \partial \lambda_k} = 0, \\
\Lambda_{jj} &= \begin{cases} \frac{\partial^2 \mathcal{L}}{\partial \lambda_j^2} = 0 \quad \text{if} \quad d^{(i)}_{j,n+1} \le d_{j,n} \quad \text{or} \quad \lambda^{(i)}_{j,n+1} \ge 0, \\ 1 \quad \text{otherwise,} \end{cases} \\
\Lambda_{jk} &= \frac{\partial^2 \mathcal{L}}{\partial \lambda_j \partial \lambda_k} = 0.
\end{aligned}
\tag{B.3}
$$

In Eqs. (21), (B.2) and (B.3), i represents the ith Newton–Raphson iteration in the same time step.

The Lagrangian multipliers are activated in two cases: (i) when $\lambda^{(i)}_{j,n+1}$ is positive, thus the damage would be negative without the enforcing equations, or (ii) when the predicted damage increment is negative. Finally, all Lagrangian equations are suppressed on nodes where the phase-field is prescribed as boundary conditions: $\bar{d}_j \rightarrow A_{jj} = 0, \Lambda_{jj} = 1$.

Appendix C. Recommendations for the use of the UEL/UMAT

With the gradual increase in options, the FORTRAN subroutine is getting more complex. Therefore, the conversion of an already existing Abaqus model to a phase-field simulation requires more attention.

Similar to the earlier implementations [14,16], the phase-field and the displacement elements are arranged in a parallel structure. The difference is that we only need to define two layers in the present case instead of three. First, the phase-field UEL then the displacement UMAT elements are created using the same nodes. The phase-field is set as the 6th, and the Lagrange multipliers as the 7th degree of freedom.

The Lagrange multipliers are added at the element level in each phase-field UEL, wherefore, the subroutine needs to have access to the connectivity table of the elements. More precisely, a separate file should be created containing the unique node numbers for each element. The present paper contains a supplementary MATLAB script that executes the conversion from an input file generated by Abaqus. Essentially it needs to identify the elements involved in the phase-field calculation using sets named with the prefix "uel_". All nodes where boundary conditions in the phase-field are defined get a value of −1. Thus, for example, for a finite element mesh with two elements defined at nodes (1, 2, 4, 3) and (3, 4, 6, 5), and with a Dirichlet boundary condition at node 6, the following connectivity file is created:

```
1, 2, 4, 3
0, 0, -1, 5
```

This file should be named the same as the job but with an extension of "_connec.dat". It is recommended to turn off extrapolation between time steps. Furthermore, due to the iterative convergence in the Lagrange multipliers when the phase-field increases significantly, the integration control is slightly modified by increasing the following parameters:

- the number of equilibrium iterations after which the convergence check is carried out to determine whether the residuals are increasing in two consecutive iterations (I_0);
- the number of consecutive equilibrium iterations at which the logarithmic rate the of convergence check begins (I_R);
- the upper limit for the number of consecutive equilibrium iterations (I_C);
- the maximum number of attempts allowed for an increment (I_A).

An example model is included as Supplementary Material.

Appendix D. Computational efficiency

The section discusses the additional computational cost added by the Lagrange multipliers. There are three main differences, which might increase the time compared to the already existing energy threshold solutions: (i) the additional iterations needed to obtain convergence; (ii) the added degrees of freedom; and (iii) the longer elastic deformation stage due to the higher strength. The latter reason originates from the difference in normalization, as explained in Fig. 12. Therefore, we will focus on only the first two causes in this section. In order to exclude the effect of the automatic time step control and focus on the effect of the Lagrange multipliers, the time steps were fixed in advance.

First, the number of iterations is analyzed using the single edge notched specimen. The geometry and the simulation details can be found in Ref. [14]. The only difference is the first stage, where we define the Dirichlet boundary condition in the phase-field to create the initial crack. This step is divided into four steps. In the first two, the boundary condition is applied. The rest of the steps are used to pass the information securely to the displacement element. The mesh size was also reduced to $l_c/5$ around the propagation zone.

Fig. D.19(a) shows the number of iterations necessary to obtain equilibrium. It is visible that the Newton–Raphson method needs a significant number of iterations for the initial Dirichlet boundary because the multipliers have to be deactivated (as shown in Fig. 3). For the AT1 solution, this usually takes $2l_c/h$ iterations in total. In this case, this number is around 12, which is close to the prediction. However, for the AT2, as the damage value is larger than 0 everywhere, it took a staggering amount of 25 in total.

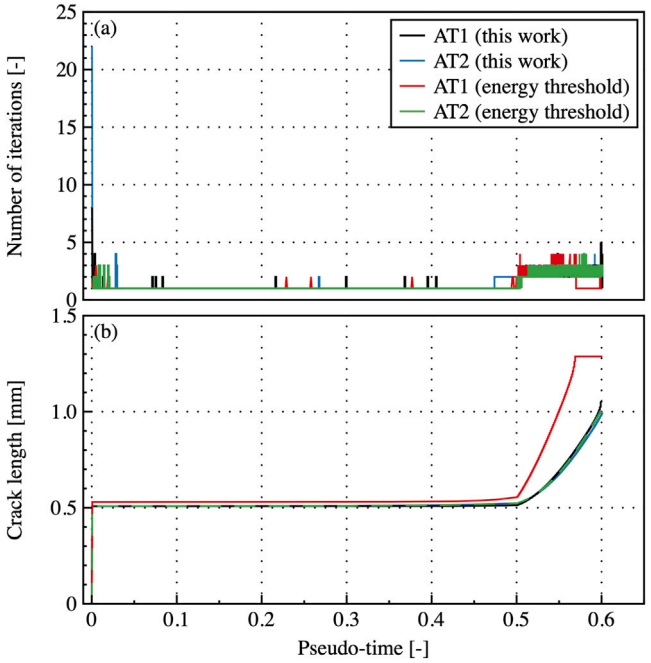

Fig. D.19. (a) Number of iterations necessary to obtain equilibrium for the different methods. (b) Crack length as a function of pseudo-time.

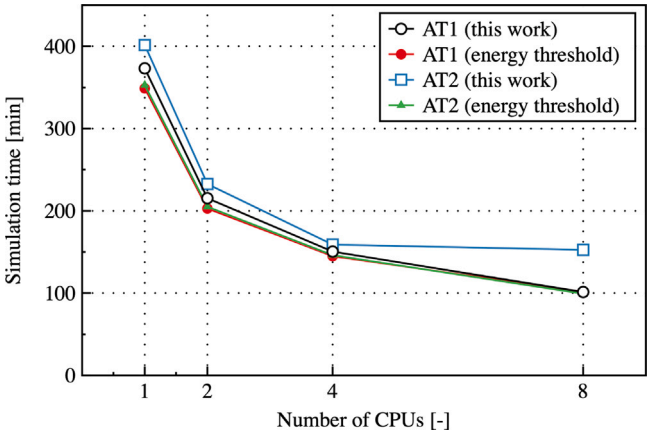

Fig. D.20. Computation time as a function of the number of CPUs for the different implementations.

While with the history variable, this takes only one iteration independently of the geometric function. However, it is worth noting that the crack length obtained using the AT1 energy threshold solution is significantly higher therefore the sample breaks at a much lower force. Furthermore, it is shown in Fig. D.19(b) that when the unstable crack propagation starts, the number of iterations is almost equal independently of the method chosen. This signifies that the change in the Lagrange multiplier is small. Thus, the crack propagates only a few element lengths. Therefore, the number of iterations necessary to obtain the mechanical equilibrium (due to the nonlinearity in the asymmetric energy degradation) is more significant.

The dynamic case presented in Section 4.3 was used to compare the time necessary to execute a simulation. The mesh was structured, and its size was increased to 0.25 mm. The larger elements allowed us to use $2 \cdot 10^{-8}$ s for the fixed time increment. The results are presented in Fig. D.20 as a function of the number of CPUs. The simulations were done using an Intel(R) Xeon(R) Gold 5220R CPU (2.20 GHz).

It is visible that using the Lagrange multiplier technique for the AT2 geometric function is not useful and increases the computation cost significantly without any major advantage. However, for the AT1 model, this increment for a single core is only 7%, and vanishes at 8 CPUs. In our experience, the computation cost in Abaqus can be attributed to two major variables: (i) the definition of the residues and stiffness matrices; (ii) compilation and solution of the linear equation system. Thanks to the efficiency of Abaqus, the second is usually significantly inferior compared to the first one. The Lagrange multipliers are added in only 30 lines

out of around 1500. Therefore, the cost in the FORTRAN subroutine is negligible. Additionally, in the AT1 threshold solution, the damage is larger than 0 in the whole model, which increases the necessary time to treat their convergence. While, in the present model, this is mostly enforced by the Lagrange multipliers.

Appendix E. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.engfracmech.2022.108390.

References

[1] Van der Waals JD. On the continuity of the gaseous and liquid states [Ph.D. thesis], Leiden; 1873.
[2] Ginzburg VL, Landau LD. On the theory of superconductivity. J Exp Theor Phys 1950;20:1064.
[3] Cahn JW, Hilliard JE. Free energy of a nonuniform system. I. Interfacial free energy. J Chem Phys 1958;28(2):258–67.
[4] Boettinger WJ, Warren JA, Beckermann C, Karma A. Phase-field simulation of solidification. Annu Rev Mater Res 2002;32(1):163–94.
[5] Mumford D, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. Comm Pure Appl Math 1989;42(5):577–685.
[6] Bourdin B, Francfort GA, Marigo J-J. Numerical experiments in revisited brittle fracture. J Mech Phys Solids 2000;48(4):797–826.
[7] Griffith AA. The phenomena of rupture and flow in solids. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 1921;221(582–593):163–98.
[8] Griffith A. The theory of rupture. In: First int. cong. appl. mech. 1924, p. 55–63.
[9] Francfort GA, Marigo J-J. Revisiting brittle fracture as an energy minimization problem. J Mech Phys Solids 1998;46(8):1319–42.
[10] Ambrosio L, Tortorelli VM. Approximation of functional depending on jumps by elliptic functional via t-convergence. Comm Pure Appl Math 1990;43(8):999–1036.
[11] Kachanov LM. Rupture time under creep conditions. Izvestia Academii Nauk Sssr Otdelenie Tekhnicheskich Nauk; 1958.
[12] Pham K, Amor H, Marigo J-J, Maurini C. Gradient damage models and their use to approximate brittle fracture. Int J Damage Mech 2011;20(4):618–52.
[13] Liu G, Li Q, Msekh MA, Zuo Z. Abaqus implementation of monolithic and staggered schemes for quasi-static and dynamic fracture phase-field model. Comput Mater Sci 2016;121:35–47.
[14] Molnár G, Gravouil A. 2D and 3D abaqus implementation of a robust staggered phase-field solution for modeling brittle fracture. Finite Elem Anal Des 2017;130:27–38.
[15] Fang J, Wu C, Rabczuk T, Wu C, Ma C, Sun G, et al. Phase field fracture in elasto-plastic solids: Abaqus implementation and case studies. Theor Appl Fract Mech 2019;103:102252.
[16] Molnár G, Gravouil A, Seghir R, Réthoré J. An open-source abaqus implementation of the phase-field method to study the effect of plasticity on the instantaneous fracture toughness in dynamic crack propagation. Comput Methods Appl Mech Engrg 2020;365:113004.
[17] Azinpour E, de Sa JC, dos Santos AD. Micromechanically-motivated phase field approach to ductile fracture. Int J Damage Mech 2021;30(1):46–76.
[18] Zhang H, Qiao P. A coupled peridynamic strength and fracture criterion for openhole failure analysis of plates under tensile load. Eng Fract Mech 2018;204:103–18.
[19] Mandal TK, Nguyen VP, Wu J-Y. Length scale and mesh bias sensitivity of phase-field models for brittle and cohesive fracture. Eng Fract Mech 2019;217:106532.
[20] Wu J-Y, Huang Y. Comprehensive implementations of phase-field damage models in Abaqus. Theor Appl Fract Mech 2020;106:102440.
[21] Cui C, Ma R, Martínez-Pañeda E. A phase field formulation for dissolution-driven stress corrosion cracking. J Mech Phys Solids 2021;147:104254.
[22] Martínez-Pañeda E, Golahmar A, Niordson CF. A phase field formulation for hydrogen assisted cracking. Comput Methods Appl Mech Engrg 2018;342:742–61.
[23] Martínez-Pañeda E, Díaz A, Wright L, Turnbull A. Generalised boundary conditions for hydrogen transport at crack tips. Corros Sci 2020;173:108698.
[24] Kristensen PK, Martínez-Pañeda E. Phase field fracture modelling using quasi-Newton methods and a new adaptive step scheme. Theor Appl Fract Mech 2020;107:102446.
[25] Wu J-Y, Huang Y, Nguyen VP. On the BFGS monolithic algorithm for the unified phase field damage theory. Comput Methods Appl Mech Engrg 2020;360:112704.
[26] Wang T, Ye X, Liu Z, Liu X, Chu D, Zhuang Z. A phase-field model of thermo-elastic coupled brittle fracture with explicit time integration. Comput Mech 2020;65(5):1305–21.
[27] Navidtehrani Y, Betegón C, Martínez-Pañeda E. A simple and robust abaqus implementation of the phase field fracture method. Appl Eng Sci 2021;6:100050.
[28] Navidtehrani Y, Betegón C, Martínez-Pañeda E. A unified abaqus implementation of the phase field fracture method using only a user material subroutine. Materials 2021;14(8).
[29] Bourdin B, Marigo J-J, Maurini C, Sicsic P. Morphogenesis and propagation of complex cracks induced by thermal shocks. Phys Rev Lett 2014;112:014301.
[30] Wu J-Y. Robust numerical implementation of non-standard phase-field damage models for failure in solids. Comput Methods Appl Mech Engrg 2018;340:767–97.
[31] Lu Y, Helfer T, Bary B, Fandeur O. An efficient and robust staggered algorithm applied to the quasi-static description of brittle fracture by a phase-field approach. Comput Methods Appl Mech Engrg 2020;370:113218.
[32] Mandal TK, Gupta A, Nguyen VP, C. R, de Vaucorbeil A. A length scale insensitive phase field model for brittle fracture of hyperelastic solids. Eng Fract Mech 2020;236:107196.
[33] Bernard P-E, Moës N, Chevaugeon N. Damage growth modeling using the thick level set (TLS) approach: Efficient discretization for quasi-static loadings. Comput Methods Appl Mech Engrg 2012;233:11–27.
[34] Miehe C, Welschinger F, Hofacker M. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. Internat J Numer Methods Engrg 2010;83(10):1273–311.
[35] Chaboche JL, Feyel F, Monerie Y. Interface debonding models: A viscous regularization with a limited rate dependency. Int J Solids Struct 2001;38(18):3127–60.
[36] Lorentz E, Benallal A. Gradient constitutive relations: Numerical aspects and application to gradient damage. Comput Methods Appl Mech Engrg 2005;194(50–52):5191–220.
[37] Wick T. Modified Newton methods for solving fully monolithic phase-field quasi-static brittle fracture propagation. Comput Methods Appl Mech Engrg 2017;325:577–611.
[38] Miehe C, Hofacker M, Welschinger F. A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits. Comput Methods Appl Mech Engrg 2010;199(45–48):2765–78.
[39] Wheeler MF, Wick T, Wollner W. An augmented-Lagrangian method for the phase-field approach for pressurized fractures. Comput Methods Appl Mech Engrg 2014;271:69–85.

[40] Geelen RJM, Liu Y, Hu T, Tupek MR, Dolbow JE. A phase-field formulation for dynamic cohesive fracture. Comput Methods Appl Mech Engrg 2019;348:680–711.
[41] Papazafeiropoulos G, Muñiz Calvente M, Martínez-Pañeda E. Abaqus2Matlab: A suitable tool for finite element post-processing. Adv Eng Softw 2017;105:9–16.
[42] Hilber HM, Hughes TJR, Taylor RL. Improved numerical dissipation for time integration algorithms in structural dynamics. Earthq Eng Struct Dyn 1977;5(3):283–92.
[43] Kristensen PK, Niordson CF, Martínez-Pañeda E. An assessment of phase field fracture: Crack initiation and growth. Phil Trans R Soc A 2021;379(2203):20210021.
[44] Molnár G, Doitrand A, Estevez R, Gravouil A. Toughness or strength? Regularization in phase-field fracture explained by the coupled criterion. Theor Appl Fract Mech 2020;109:102736.
[45] Molnár G, Ganster P, Tanguy A, Barthel E, Kermouche G. Densification dependent yield criteria for sodium silicate glasses – An atomistic simulation approach. Acta Mater 2016;111:129–37.
[46] Rankine WJM. II. On the stability of loose earth. Philos Trans R Soc Lond 1857;(147):9–27.
[47] Gravouil A, Pierres E, Baietto MC. Stabilized global–local X-FEM for 3D non-planar frictional crack using relevant meshes. Internat J Numer Methods Engrg 2011;88(13):1449–75.
[48] Chen Y. Modeling of ductile fracture using local approach: Reliable simulation of crack extension [Ph.D. thesis], Université Paris sciences et lettres; 2019.
[49] Seleš K, Lesičar T, Tonković Z, Sorić J. A residual control staggered solution scheme for the phase-field modeling of brittle fracture. Eng Fract Mech 2019;205:370–86.
[50] Bourdin B, Francfort GA, Marigo J-J. The variational approach to fracture. Springer Netherlands; 2008.
[51] Linse T, Hennig P, Kästner M, de Borst R. A convergence study of phase-field models for brittle fracture. Eng Fract Mech 2017;184:307–18.
[52] Tanné E, Li T, Bourdin B, Marigo J-J, Maurini C. Crack nucleation in variational phase-field models of brittle fracture. J Mech Phys Solids 2018;110:80–99.
[53] Bažant ZP. Size effect on structural strength: A review. Arch Appl Mech 1999;69(9–10):703–25.
[54] Lazarus V. Three-dimensional linear elastic fracture mechanics: From theory to practice [Ph.D. thesis], Université Pierre et Marie Curie-Paris VI; 2010.
[55] Lazarus V, Buchholz F-G, Fulland M, Wiebesiek J. Comparison of predictions by mode II or mode III criteria on crack front twisting in three or four point bending experiments. Int J Fract 2008;153(2):141–51.
[56] Pham KH, Ravi-Chandar K. Further examination of the criterion for crack initiation under mixed-mode I + III loading. Int J Fract 2014;189(2):121–38.
[57] Pham KH, Ravi-Chandar K. On the growth of cracks under mixed-mode I + III loading. Int J Fract 2016;199(1):105–34.
[58] Pons AJ, Karma A. Helical crack-front instability in mixed-mode fracture. Nature 2010;464(7285):85–9.
[59] Chen C-H, Cambonie T, Lazarus V, Nicoli M, Pons AJ, Karma A. Crack front segmentation and facet coarsening in mixed-mode fracture. Phys Rev Lett 2015;115(26):265503.
[60] Knauss WG. An observation of crack propagation in anti-plane shear. Int J Fract Mech 1970;6(2):183–7.
[61] Lazarus V, Leblond J-B, Mouchrif S-E. Crack front rotation and segmentation in mixed mode I + III or I + II + III. Part II: Comparison with experiments. J Mech Phys Solids 2001;49(7):1421–43.
[62] Lazarus V, Leblond J-B, Mouchrif S-E. Crack front rotation and segmentation in mixed mode I + III or I + II + III. Part I: Calculation of stress intensity factors. J Mech Phys Solids 2001;49(7):1399–420.
[63] Leblond J-B, Karma A, Lazarus V. Theoretical analysis of crack front instability in mode I + III. J Mech Phys Solids 2011;59(9):1872–87.
[64] Pham KH, Ravi-Chandar K. The formation and growth of echelon cracks in brittle materials. Int J Fract 2017;206(2):229–44.
[65] Jin Z, Sun C-T. Fracture mechanics. Elsevier; 2011.
[66] Gao H. A theory of local limiting speed in dynamic fracture. J Mech Phys Solids 1996;44(9):1453–74.
[67] Kalthoff JF, Winkler S. Failure mode transition at high rates of shear loading. In: DGM informationsgesellschaft mbH, impact loading and dynamic behavior of materials, vol. 1, 1988, p. 185–95.
[68] Klinsmann M, Rosato D, Kamlah M, McMeeking R-M. An assessment of the phase field formulation for crack growth. Comput Methods Appl Mech Engrg 2015;294:313–30.
[69] Kumar A, Bourdin B, Francfort GA, Lopez-Pamies O. Revisiting nucleation in the phase-field approach to brittle fracture. J Mech Phys Solids 2020;142:104027.
[70] Luo J, Wang J, Bitzek E, Huang JY, Zheng H, Tong L, et al. Size-dependent brittle-to-ductile transition in silica glass nanofibers. Nano Lett 2016;16(1):105–13.