Comput. Methods Appl. Mech. Engrg. 200 (2011) 1927–1944

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

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

Gradient damage models: Toward full-scale computations

E. Lorentz [a,b,*], V. Godard [b]

[a] Laboratoire de Mécanique des Structures Industrielles Durables, UMR 2832 CNRS/ EDF, 1 av. Général de Gaulle, 92141 Clamart cedex, France
[b] EDF Research & Development, 1 av. Général de Gaulle, 92141 Clamart cedex, France

ARTICLE INFO

Article history:
Received 17 June 2009
Received in revised form 31 January 2010
Accepted 23 June 2010
Available online 13 July 2010

Keywords:
Damage
Strain-softening
Gradient constitutive laws
Variational formulation
Augmented lagrangian

ABSTRACT

The description of localisation and structural failure through local damage constitutive laws leads to ill-posed problems. To retrieve a well-posed problem, the material state is described through the gradients of damage in addition to the strain and the damage. Under some assumptions, it is shown that these gradient laws can be cast into a variational formulation. Its numerical treatment is reduced to the design of a mixed finite element so that algorithms such as Newton's method and path-following techniques are still available. Possible convergence difficulties related to the reduction of the damage band width and the appearance of micro instabilities are removed thanks to the design of a new class of brittle laws. Finally, numerical applications show the convergence with mesh refinement, the compatibility with coarser approaches such as Griffith's theory and cohesive crack models, the robustness in case of unstable multi-cracking, the applicability to full-scale 3D situations and the physical relevance compared to experimental data.

1. Introduction

Continuum Damage Mechanics seems an appealing approach to describe the different steps of structural failure within a single framework [1]: damage initiation, micro-crack growth, coalescence and macro-crack propagation up to the ultimate failure of the structure. However, local damage constitutive relations are known to lead to ill-posed problems caused by strain-softening, see [2] for a review, resulting in strongly mesh-dependent computations [3]. On a physical ground, damage localises in narrow bands so that the length scales of macroscopic fields and of material microstructures are no more separated, hence hindering the validity of macroscopic local laws. Nonlocal constitutive relations have been introduced in order to take into account and somewhat control the high gradients of damage. Despite some attempts to derive the nonlocal terms through homogenisation, see [4] or [5] for instance, they are mostly introduced on the basis of phenomenological or pragmatic considerations. This leads to several classes of formulations, among which:

- the regularisation of state variables, as analysed in [6], with namely the well-known "integral" nonlocal model [7] and the "implicit gradient" model [8];
- the introduction of higher order gradients of the displacement field, which encompasses in particular strain gradient models [9,10];
- the introduction of gradients of internal variable fields [11,12] or new kinematical degrees of freedom [13,14], both approaches being equivalent in quasi-static regime [15].

Following earlier work in [15,16], we focus our attention on brittle damage regularised by the introduction of damage gradient. The choice is motivated by some attractive properties: (i) the meaning of the classical stress tensor is preserved regarding the equilibrium equations, which simplifies the analysis of computational results by non-expert engineers; (ii) the regularised quantity is a field that totally controls the localisation process (no localisation can occur without high gradient of this particular field) while limiting the increase of the number of unknowns thanks to the scalar nature of damage; (iii) the damage variable is also a bounded field (on the contrary of the strain field) which should preclude issues related to a non controlled increase of the localisation band width [17]. Despite these properties, such a nonlocal formulation has seemingly received less attention in the literature than others. We think this can be attributed to the fact that the consistency condition becomes nonlocal, hence raising special numerical issues. Thus, both alternatives of the consistency condition (loading/unloading) have been taken into account in [18] thanks to the introduction of Lagrange multipliers, in [19] by means of a mixed dual formulation and in [20] relying on a relaxation by a penalty method; but the applications stayed limited to small scale problems. In addition, it has been observed in [12,19] that the localisation band tends to narrow with increasing damage, hence ending up with mesh-dependency, a serious drawback of the approach. However, significant success have been obtained on 3D geometries

* Corresponding author. Tel.: +33 1 47 65 35 81; fax: +33 1 47 65 54 28.
E-mail address: Eric.Lorentz@edf.fr (E. Lorentz).

0045-7825/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2010.06.025

in the context of Regularised Fracture Mechanics [21] although the formulation is close to gradient damage. The main differences lie in the lack of an initial damage threshold, the presence of a residual stiffness (not a true crack, resulting in possibly high stress values) and in the relaxation of the irreversibility condition (damage is not enforced to be irreversible). These features seem questionable in the context of Damage Mechanics and deserve additional attention in order to estimate their influence.

Nevertheless encouraged by the promising results of Regularised Fracture Mechanics, we aim here at defining a gradient damage formulation that preserves the features of the local constitutive relation (namely a damage threshold, zero ultimate stiffness and irreversibility) while trying to answer the questions of mesh dependency, computational efficiency and limited intrusion in finite element codes so as to preserve the compatibility with pre-existing solution algorithms and path-following methods. This goal is achieved in three steps. First, following [16], the nonlocal constitutive relation receives a global expression in Section 2 by taking advantage of the variational properties of generalised standard materials [22,23]. Then, in Section 3, the highly nonlinear character is dealt thanks to a decomposition-coordination technique expressed on the continuous problem, before spatial discretisation. It results in the derivation of a mixed finite element, the degrees of freedom of which are the displacements, the damage and a scalar Lagrange multiplier. A symmetric consistent tangent matrix is available, hence enabling computational efficiency. At last, the choice of the damage law and especially its softening part is analysed in Section 4 by means of a 1D problem in order to get salient features with straightforward consequences on numerical robustness (convergence) and mesh-dependency. The analysis is led on the basis of a simple class of isotropic brittle law in order to focus on the difficulties related to nonlocality; in particular, distinction between tension and compression, crack closure, anisotropy and irreversible strain are not taken into account at this stage. The capabilities of the model are demonstrated through several numerical investigations in Section 5 and confronted to experimental observations in Section 6.

2. Gradient constitutive relations

2.1. The formalism of generalised standard materials at the scale of the structure

In order to benefit from variational properties for the constitutive relations, we restrict our attention to generalised standard materials, even though the resulting discrete formulation could be used beyond this framework. Moreover, the analysis is led in the context of isotropic brittle damage. In that case, the material state may be defined by the strain tensor $\boldsymbol{\varepsilon}$ and a scalar damage variable a. The damage gradient ∇a is also introduced in the constitutive law in order to control the localisation by coupling the behaviour of neighbour material points. But as explained in [15,16], the value and the gradient of the damage field a cannot be considered as independent variables since their evolution is constrained by a nonlocal relation (being the value and the gradient of a single field). Therefore, the framework of generalised standard materials is extended to the structure scale, as done in [23]: the free energy $\mathcal{F}$ and the dissipation potential $\mathcal{D}$ of the structure are defined as functionals of the strain and the damage fields. We propose the following expressions:

$$\mathcal{F}(\boldsymbol{\varepsilon}, a) = \int_\Omega \Phi(\boldsymbol{\varepsilon}, a) d\Omega + \int_\Omega \frac{c}{2} (\nabla a)^2 d\Omega, \tag{1}$$

$$\mathcal{D}(\dot{a}) = \int_\Omega \Psi(\dot{a}) d\Omega, \tag{2}$$

where Ω denotes the body domain and a dot stands for time differentiation. Φ is the local free energy; it should be a convex function with respect to $\boldsymbol{\varepsilon}$ and a separately. Ψ is the local dissipation potential which should also be convex and minimal in 0. And $c > 0$ is a parameter that controls indirectly the localisation band width through the influence of damage gradients on the free energy.

At this stage, the global constitutive relations are derived from these potentials:

$$\boldsymbol{\sigma} = \frac{\partial \mathcal{F}}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}, a), \tag{3}$$

$$\mathcal{Y} = -\frac{\partial \mathcal{F}}{\partial a}(\boldsymbol{\varepsilon}, a), \tag{4}$$

$$\mathcal{Y} \in \partial \mathcal{D}(\dot{a}). \tag{5}$$

$\boldsymbol{\sigma}$ and $\mathcal{Y}$ denote the stress and the driving force associated to damage. Mathematically, they are linear forms operating respectively on the strain and damage fields. $\partial \mathcal{D}$ is the sub-gradient of $\mathcal{D}$, an extension of the notion of derivative for convex non-smooth functions. It is necessary in the case of rate-independent constitutive laws where the dissipation potential is positive homogeneous of degree 1 with respect to $\dot{a}$, hence generally not differentiable. For the sake of completeness, its definition is recalled here (where the brackets stand for the duality product), see for instance [24] for complementary propositions and related properties:

$$\mathcal{Y} \in \partial \mathcal{D}(\dot{a}) \Longleftrightarrow \quad \forall \delta a \ \mathcal{D}(\delta a) \geqslant \mathcal{D}(\dot{a}) + \langle \mathcal{Y} | \delta a - \dot{a} \rangle. \tag{6}$$

Even though the constitutive relations are totally defined through (3)–(5), this stays nevertheless quite formal and does not provide practical expressions. We will show in the next parts how these relations can be expressed first in a pointwise manner, then as a minimisation problem.

2.2. Pointwise interpretation of the gradient damage model

In the stress–strain relation (3), the strain field is involved only in a pointwise manner. Therefore, application of the linear form $\boldsymbol{\sigma}$ to any virtual strain field $\delta\boldsymbol{\varepsilon}$ simply leads to the usual expression of the internal virtual work:

$$\langle \boldsymbol{\sigma} | \delta \boldsymbol{\varepsilon} \rangle = \left\langle \frac{\partial \mathcal{F}}{\partial \boldsymbol{\varepsilon}} \middle| \delta \boldsymbol{\varepsilon} \right\rangle = \int_\Omega \frac{\partial \Phi}{\partial \boldsymbol{\varepsilon}} : \delta \boldsymbol{\varepsilon} \, dx. \tag{7}$$

Thus, the equilibrium equations and the stress retain their usual local meaning, where the stress field is given by:

$$\boldsymbol{\sigma} = \frac{\partial \Phi}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}, a). \tag{8}$$

Note that for the sake of simplicity, the notations are chosen identical for linear forms, corresponding fields (when possible through the theorem of representation) and pointwise values (explicit reference to the position $\mathbf{x}$ is omitted).

In the combination of the state Eq. (4) and the evolution Eq. (5), the damage field is involved not only through its value but also through its gradient. On the contrary to what has been done for the stress, it is not possible to identify the driving force $\mathcal{Y}$ (linear form) with a single field Y. An application of the rules of the calculus of variations leads to the following expression where δa denotes a virtual damage field:

$$\langle \mathcal{Y} | \delta a \rangle = \int_\Omega -\frac{\partial \Phi}{\partial a}(\boldsymbol{\varepsilon}, a) \delta a \, d\Omega + \int_\Omega -c \nabla a \cdot \nabla \delta a \, d\Omega. \tag{9}$$

It is then possible to define a couple field $(Y, \mathbf{Y})$ of (scalar and vector) driving forces:

$$Y = -\frac{\partial \Phi}{\partial a}(\boldsymbol{\varepsilon}, a); \quad \mathbf{Y} = -c \nabla a. \tag{10}$$

By identification with (9), the relation with the linear form $\mathcal{Y}$ is immediate (and results actually from the theorem of representation):

$$\langle \mathcal{Y} | \delta a \rangle = \int_\Omega (Y \delta a + \mathbf{Y} \nabla \delta a) dx. \tag{11}$$

An integration by parts leads to the equivalent expression of the driving force $\mathcal{Y}$:

$$\langle \mathcal{Y} | \delta a \rangle = \int_\Omega (Y - \nabla \cdot \mathbf{Y}) \delta a \, dx + \int_{\partial\Omega} \mathbf{Y} \cdot \mathbf{n} \delta a \, ds - \int_\Gamma [\![\mathbf{Y}]\!] \cdot \boldsymbol{v} \delta a \, ds, \tag{12}$$

where $\partial\Omega$ denotes the boundary of the body domain Ω, $\mathbf{n}$ its outer normal, Γ a potential surface of discontinuity, $\boldsymbol{v}$ its normal and $[\![\cdot]\!]$ the discontinuity of a quantity across Γ computed along v. Besides, the dissipation potential (2) depends on the damage rate only through its value (and not its gradient) and it does not involve surface terms. Following the arguments in [25], both properties enable a simple interpretation of the differential inclusion (5) with respect to each term of (12), yielding respectively a pointwise relation, a boundary condition and an interface condition on Γ:

$$Y - \nabla \cdot \mathbf{Y} \in \partial \Psi(\dot{a}) \quad \text{in } \Omega, \tag{13}$$

$$\mathbf{Y} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega, \tag{14}$$

$$[\![\mathbf{Y}]\!] \cdot \boldsymbol{v} = 0 \quad \text{across } \Gamma. \tag{15}$$

Actually, the boundary and the interface conditions are a priori less restrictive than (14) and (15) when the dissipation potential involves indicator functions (thus resulting in surface terms). But it could be shown that the possible additional solutions are not compatible with the evolution of $\mathbf{Y}$ in time so that we can indeed restrict our attention to (13)–(15).

Moreover, the damage field is subjected to a regularity condition so that the free energy $\mathcal{F}$ be finite. Because of the term in $(\nabla a)^2$, a natural choice is $a \in H^1(\Omega)$. In particular, it implies that:

$$[\![a]\!] = 0. \tag{16}$$

Finally, the Eqs. (8) and (10) provide pointwise state laws. The differential inclusion (13) characterises the evolution of damage. And (14)–(16) express the interface and boundary conditions. To go one step further in this pointwise interpretation, the differential inclusion should be made more explicit. In the case of rate-dependent evolution laws for which the local dissipation potential Ψ is differentiable, the subgradient simply reduces to the derivative $\partial\Psi(\dot{a}) = \{\Psi'(\dot{a})\}$. In the case of rate-independent evolution law, the dissipation potential is positive homogeneous of degree 1. The differential inclusion is then equivalent to a flow rule and a consistency condition; there is a convex function f so that:

Box 1. Application to a class of brittle gradient damage models.

Local potentials

Free Helmholtz' energy $\Phi(\boldsymbol{\varepsilon}, a) = \frac{1}{2} \mathrm{A}(a) \boldsymbol{\varepsilon} : \mathbf{E} : \boldsymbol{\varepsilon} + \omega(a) - ka$ with $0 \leqslant a \leqslant 1$

Dissipation potential $\Psi(\dot{a}) = k\dot{a} + \mathrm{I}_{\mathbb{R}^+}(\dot{a})$

Pointwise interpretation of the gradient law

Stress tensor $\boldsymbol{\sigma} = \mathrm{A}(a)\mathbf{E}:\boldsymbol{\varepsilon}$

Damage driving force $Y = -\frac{1}{2} A'(a) \boldsymbol{\varepsilon} : \mathbf{E} : \boldsymbol{\varepsilon} - \omega'(a) + k$

Consistency condition $Y + c\nabla^2 a - k \leqslant 0; \quad \dot{a} \geqslant 0; \quad \dot{a}[Y + c\nabla^2 a - k] = 0$

Boundary condition $\nabla a \cdot \mathbf{n} = 0$ on $\partial\Omega$

Interface conditions $[\![a]\!] = 0; \; [\![c \nabla a]\!] \cdot \boldsymbol{v} = 0$

Time-discretisation

$(\mathbf{u}^{n+1}, a^{n+1}, \lambda^{n+1}, \alpha^{n+1})$ saddle-point of the following augmented Lagrangian:

$$\mathcal{L}_r(\mathbf{u}, a, \lambda, \alpha) = \int_\Omega [\Phi(\boldsymbol{\varepsilon}(\mathbf{u}), \alpha) + \Psi(a - a^n)] - \mathcal{W}_{\text{ext}}(\mathbf{u}) + \int_\Omega \frac{c}{2} (\nabla a)^2 + \int_\Omega \frac{r}{2} (a - \alpha)^2 + \int_\Omega \lambda (a - \alpha)$$

Pointwise constitutive relation

Damage evolution $\alpha_g^{n+1} = \hat{\alpha}\left(\boldsymbol{\varepsilon}_g^{n+1}, Z_g^{n+1}\right)$ where $Z_g^{n+1} = \mathbf{N}_g^{\lambda} \mathbf{L}^{n+1} + r \mathbf{N}_g^{a} \mathbf{A}^{n+1}$

Elastic prediction if $Y(\boldsymbol{\varepsilon}, \alpha^n) + Z - r\,\alpha^n - k \leqslant 0$ then $\hat{\alpha}(\boldsymbol{\varepsilon}, Z) = \alpha^n$

Damage correction else $\hat{\alpha}(\boldsymbol{\varepsilon}, Z) = \alpha \geqslant \alpha^n$ so that $Y(\boldsymbol{\varepsilon}, \alpha) + Z - r\alpha - k = 0$

Stress–strain relation $\boldsymbol{\sigma}_g^{n+1} = \hat{\boldsymbol{\sigma}}\left(\boldsymbol{\varepsilon}_g^{n+1}, Z_g^{n+1}\right) = A\left(\hat{a}\left(\boldsymbol{\varepsilon}_g^{n+1}, Z_g^{n+1}\right)\right) \mathbf{E} : \boldsymbol{\varepsilon}_g^{n+1}$

Residual nodal vectors and consistent tangent matrix

$$\begin{cases} \mathbf{R}^u = \mathbf{F}_{\text{ext}} \\ \mathbf{R}^a = \mathbf{0} \\ \mathbf{R}^\lambda = \mathbf{0} \end{cases} \quad \text{with} \quad \begin{bmatrix} \mathbf{R}^u \\ \mathbf{R}^a \\ \mathbf{R}^\lambda \end{bmatrix} = \sum_g w_g \begin{bmatrix} \mathbf{B}_g^{u^T} \boldsymbol{\sigma}_g \\ c\mathbf{B}_g^{a^T} \mathbf{B}_g^a \mathbf{A} + r\mathbf{N}_g^{a^T} \left(\mathbf{N}_g^a \mathbf{A} - \alpha_g\right) + \mathbf{N}_g^{a^T} \mathbf{N}_g^\lambda \mathbf{L} \\ \mathbf{N}_g^{\lambda^T} \left(\mathbf{N}_g^a \mathbf{A} - \alpha_g\right) \end{bmatrix}$$

$$\mathbf{K} = \sum_g w_g \begin{bmatrix} \mathbf{B}_g^{u^T} \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial \boldsymbol{\varepsilon}} \mathbf{B}_g^u & r\mathbf{B}_g^{u^T} \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial Z} \mathbf{N}_g^a & \mathbf{B}_g^{u^T} \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial Z} \mathbf{N}_g^\lambda \\ \text{sym.} & c\mathbf{B}_g^{a^T} \mathbf{B}_g^a + r\left(1 - r\frac{\partial \hat{\alpha}}{\partial Z}\right) \mathbf{N}_g^{a^T} \mathbf{N}_g^a & \left(1 - r\frac{\partial \hat{\alpha}}{\partial Z}\right) \mathbf{N}_g^{a^T} \mathbf{N}_g^\lambda \\ \text{sym.} & \text{sym.} & -\frac{\partial \hat{\alpha}}{\partial Z} \mathbf{N}_g^{\lambda^T} \mathbf{N}_g^\lambda \end{bmatrix}$$

$$y \in \partial\Psi(\dot{a}) \Longleftrightarrow \begin{cases} \dot{a} = \mu \dfrac{df}{dy}(y), \\ f(y) \leqslant 0; \quad \mu \geqslant 0; \quad \mu f(y) = 0. \end{cases} \tag{17}$$

For instance, consider the following choice of dissipation potential:

$$\Psi(\dot{a}) = k\dot{a} + I_{\mathbb{R}^+}(\dot{a}) \quad \text{with} \quad I_{\mathbb{R}^+}(\dot{a}) = \begin{cases} 0 & \text{if } \dot{a} \geqslant 0, \\ +\infty & \text{if } \dot{a} < 0, \end{cases} \tag{18}$$

where $k \geqslant 0$. The indicator function $I_{\mathbb{R}^+}$ is introduced in order to enforce the irreversibility by prohibiting negative damage rate. Note that Ψ is indeed convex, minimal in 0 and positive homogeneous of degree 1. A classical application of convex analysis results shows that the threshold function simply reads:

$$f(y) = y - k. \tag{19}$$

The application to the evolution Eq. (13) along with the definition (10) of $\mathbf{Y}$ give:

$$f(Y - \nabla \cdot \mathbf{Y}) = Y + c\nabla^2 a - k. \tag{20}$$

The gradient model defined in (1)–(5) appears equivalent to the introduction of a Laplacian of the damage field inside the threshold function of the local model. An application of these results is given in box 1 for a class of brittle constitutive laws.

2.3. Time-step integration: a minimum energy problem

As the equilibrium equations remain unchanged, thanks to (7), "only" the integration of the gradient constitutive equations will deserve special attention. Even though a pointwise interpretation has been given to the gradient damage model (1)–(5), it does not help the numerical treatment because of the differential nature of the consistency condition expressed in (17) and (20). It seems preferable to take advantage of the formulation (4), (5) of the constitutive relation.

The integration with respect to time is performed by means of an implicit Euler scheme, motivated (i) by the stability provided by an exact satisfaction of the consistency condition at the end of each time-step; (ii) by the fact that the poor regularity of the damage rate field precludes explicit methods, on the contrary to implicit ones for which existence results have been demonstrated [26]. More precisely, consider a given time-step $[t^n, t^{n+1}]$, where q^n denotes the value of a quantity q at the beginning of the time-step, q^{n+1} its value at the end of the time-step and $\Delta q = q^{n+1} - q^n$ its increment over the time-step. The problem of time integration can be stated as follows: the strain field $\boldsymbol{\varepsilon}^n$ and the damage field a^n being known, find the damage field a^{n+1} corresponding to a given final strain field $\boldsymbol{\varepsilon}^{n+1}$. A straightforward application of the implicit Euler scheme reads:

$$-\frac{\partial \mathcal{F}}{\partial a}(\boldsymbol{\varepsilon}^{n+1}, a^{n+1}) \in \partial\mathcal{D}\left(\frac{\Delta a}{\Delta t}\right). \tag{21}$$

It is then possible to take advantage of the assumed convexity of the potentials with respect to a to express this differential inclusion as a minimum principle [24]:

$$a^{n+1} = \arg\min_{a} \left[\mathcal{F}(\boldsymbol{\varepsilon}^{n+1}, a) + \Delta t \mathcal{D}\left(\frac{a - a^n}{\Delta t}\right)\right]. \tag{22}$$

Therefore, we are naturally led to introduce the (incremental) potential energy, where the explicit reference to the initial state a^n has been omitted for the sake of simplicity:

$$\mathcal{E}(\mathbf{u}, a) = \mathcal{F}(\boldsymbol{\varepsilon}(\mathbf{u}), a) + \Delta t \mathcal{D}\left(\frac{a - a^n}{\Delta t}\right) - \mathcal{W}_{ext}(\mathbf{u}). \tag{23}$$

$\mathbf{u}$ and $\mathcal{W}_{ext}$ denote respectively the displacement field and the potential of external work. Thanks to (7) and (22), the equilibrium equations and the discretised constitutive relations are expressed as combined minima of the potential energy (with KA the set of kinematically admissible displacements):

$$\mathbf{u}^{n+1} = \arg\min_{\mathbf{u} \in KA} \mathcal{E}(\mathbf{u}, a^{n+1}); \quad a^{n+1} = \arg\min_{a} \mathcal{E}(\mathbf{u}^{n+1}, a). \tag{24}$$

3. Numerical solution

Formulating the problem as the minimisation (24) enables the use of dedicated mathematical tools. However, at least in the case of rate-independent materials, the potential energy $\mathcal{E}$ may be highly nonlinear because the dissipation potential is positive homogeneous of degree 1; it is the counterpart of embedding the consistency condition. Therefore, most of minimisation techniques that work for smooth problems are doomed to fail. Moreover, the target in terms of number of degrees of freedom is sufficiently high so that algorithms specifically dedicated to non differentiable functions may not be efficient enough. Therefore, we focus our attention on a solution method which separates the difficulties between a linear global problem (structure scale) and nonlinear local problems (finite element level) by means of a decomposition – coordination technique [27], ending up with a mixed variational setting based on an augmented Lagrangian. Note that a summary of the main results is given again in box 1 for a special class of brittle laws.

3.1. Mixed variational setting

The main idea consists in introducing an auxiliary field $\alpha \in L^2(\Omega)$ assumed to be equal to the damage field $a \in H^1(\Omega)$. The potential energy is then expressed as:

$$\widehat{\mathcal{E}}(\mathbf{u}, a, \alpha) = \int_\Omega \Pi(\boldsymbol{\varepsilon}(\mathbf{u}), \alpha) - \mathcal{W}_{\text{ext}}(\mathbf{u}) + \int_\Omega \frac{c}{2}(\nabla a)^2$$
$$\text{with} \quad \Pi(\boldsymbol{\varepsilon}, \alpha) \underset{def.}{=} \Phi(\boldsymbol{\varepsilon}, \alpha) + \Delta t \Psi\left(\frac{\alpha - \alpha^n}{\Delta t}\right). \tag{25}$$

Note that the new field is the argument of the non differentiable part Π while the former damage field contributes through its gradient. In order to gain some additional coercivity, a penalty term is added to the potential energy with $r > 0$, without influence as soon as the constraint $a = \alpha$ is fulfilled:

$$\widehat{\mathcal{E}}_r(\mathbf{u}, a, \alpha) = \int_\Omega \Pi(\boldsymbol{\varepsilon}(\mathbf{u}), \alpha) - \mathcal{W}_{\text{ext}}(\mathbf{u}) + \int_\Omega \frac{c}{2}(\nabla a)^2 + \int_\Omega \frac{r}{2}(a - \alpha)^2. \tag{26}$$

The minimisation problem (24) is then strictly equivalent to the minimisation of (26) with respect to a and α subjected to $a = \alpha$. Finally, the constraint is enforced through dualisation; the problem consists in finding the saddle-point of the following (augmented) Lagrangian:

$$\mathcal{L}_r(\mathbf{u}, a, \lambda, \alpha) = \int_\Omega \Pi(\boldsymbol{\varepsilon}(\mathbf{u}), \alpha) - \mathcal{W}_{\text{ext}}(\mathbf{u}) + \int_\Omega \frac{c}{2}(\nabla a)^2 + \int_\Omega \frac{r}{2}(a - \alpha)^2 + \int_\Omega \lambda(a - \alpha), \tag{27}$$

where $\lambda \in L^2(\Omega)$ is the Lagrange multiplier. For a given displacement field $\mathbf{u}$, the problem of minimising (26) admits a unique solution $\alpha = a \in H^1(\Omega)$ thanks to the strict convexity of Π with respect to α, its lower semi-continuity and its coercivity in $L^\infty(\Omega)$ (for bounded damage). Thanks again to the convexity, finding the saddle-point of (27) is then equivalent to the former constraint minimisation.

However, the question of the existence of a solution to the global problem, including the unknown displacement field, remains an open problem to the authors' knowledge. This is due to the lack of convexity and coercivity of Π with respect to the pair $(\boldsymbol{\varepsilon}, \alpha)$. This

difficult question is disregarded here and only the optimality conditions of a saddle-point $(\mathbf{u}, a, \lambda, \alpha)$ are considered. They read:

$$\forall \delta\mathbf{u} \in KA^0 \quad \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\delta\mathbf{u}) = \mathcal{W}'_{ext}(\delta\mathbf{u}) \quad \text{with} \quad \boldsymbol{\sigma}(\mathbf{x}) = \frac{\partial \Phi}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\mathbf{x}), \alpha(\mathbf{x})), \tag{28}$$

$$\forall \delta a \in H^1(\Omega) \quad \int_\Omega c\nabla a \cdot \nabla \delta a + \int_\Omega r(a-\alpha)\delta a + \int_\Omega \lambda \delta a = 0, \tag{29}$$

$$\forall \delta\lambda \in L^2(\Omega) \quad \int_\Omega (a-\alpha)\delta\lambda = 0, \tag{30}$$

$$\forall \mathbf{x} \in \Omega \quad \lambda(\mathbf{x}) + r(a(\mathbf{x}) - \alpha(\mathbf{x})) \in \partial_\alpha \Pi(\boldsymbol{\varepsilon}(\mathbf{x}), \alpha(\mathbf{x})). \tag{31}$$

The variational Eq. (28) corresponds to the equilibrium of the structure, with KA^0 the set of kinematically admissible virtual displacements with homogeneous Dirichlet conditions. (29) is the linear global problem with respect to a. (30) enforces the linear constraint $a = \alpha$. And finally, (31) corresponds to a local constitutive relation, with the left-hand side as a perturbation term. Indeed, it also reads:

$$\forall \mathbf{x} \in \Omega \quad \underbrace{-\frac{\partial \Phi}{\partial a}(\boldsymbol{\varepsilon}(\mathbf{x}), \alpha(\mathbf{x}))}_{\text{local driving force}} + \underbrace{\lambda(\mathbf{x}) + r(a(\mathbf{x}) - \alpha(\mathbf{x}))}_{\text{nonlocal contribution}} \in \partial \Psi(\alpha(\mathbf{x})). \tag{32}$$

It shows that the local driving force is corrected by a pointwise nonlocal contribution into the usual local constitutive relation.

The idea of relaxation through the introduction of an additional field has already been applied to gradient constitutive laws in [19], with the following differences compared to the present formulation: the authors introduced a (vector) field corresponding to the gradient of cumulated plasticity, derived a non symmetrical variational formulation in which no stabilising term was introduced and proceeded to a spatial discretisation based on low order mixed finite elements. In comparison, the introduction of the penalty term might be seen as a drawback, but it has little influence in practice. Conversely, relaxing the damage field itself results in less unknowns in 3D along with a symmetrical problem and a natural discretisation by second order displacement finite elements. Moreover, retaining the damage field as a primal variable is more convenient as soon as damage essential boundary conditions have to be prescribed; even though not done in the present work, this may be the case for cracks that propagate along a boundary, see [21]. And finally, preserving an energy formulation enables the use of stability criteria as proposed in [25].

3.2. Spatial discretisation by mixed finite elements

3.2.1. Degrees of freedom

The displacement $\mathbf{u}$, the damage field a and the Lagrange multiplier field λ are discretised by means of continuous Lagrange shape functions. In matrix format, it reads:

$$\begin{aligned} &\mathbf{u}(\mathbf{x}) = \mathbf{N}^u(\mathbf{x}) \cdot \mathbf{U}; \quad \boldsymbol{\varepsilon}(\mathbf{x}) = \mathbf{B}^u(\mathbf{x}) \cdot \mathbf{U}, \\ &a(\mathbf{x}) = \mathbf{N}^a(\mathbf{x}) \cdot \mathbf{A}; \quad \nabla a(\mathbf{x}) = \mathbf{B}^a(\mathbf{x}) \cdot \mathbf{A}, \\ &\lambda(x) = \mathbf{N}^\lambda(\mathbf{x}) \cdot \mathbf{L}, \end{aligned} \tag{33}$$

where $\mathbf{U} \in \mathbb{R}^{n^u}$, $\mathbf{A} \in \mathbb{R}^{n^a}$ and $\mathbf{L} \in \mathbb{R}^{n^\lambda}$ are the nodal vectors corresponding to the displacement, the damage and the Lagrange multiplier unknowns. The matrices $\mathbf{B}^u$ and $\mathbf{B}^a$ are obtained through spatial derivation of $\mathbf{N}^u$ and $\mathbf{N}^a$ respectively.

Following the ideas of decomposition–coordination [27], the discretisation of the additional field α is based on collocation points of co-ordinate $\mathbf{x}_g$:

$$\alpha(\mathbf{x}_g) = \alpha_g, \tag{34}$$

where α_g corresponds to the degrees of freedom characterising the field α. The set of collocation points is chosen so that it coincides with the integration points used in the quadrature rule, with weight w_g. More precisely, the quadrature rule reads for any function q:

$$\int_\Omega q(\mathbf{x}) d\Omega \simeq \sum_g w_g q_g \quad \text{with} \quad q_g = q_(\mathbf{x}_g), \tag{35}$$

where the subscript g denotes the evaluation of a function at $\mathbf{x}_g$.

3.2.2. Non linear pointwise problems

Thanks to the collocation discretisation (34), the relation (31) which condensates the high nonlinearity of Π retains a pointwise interpretation, close to the local constitutive law:

$$\forall g \quad Z_g - r\alpha_g \in \partial_\alpha \Pi(\boldsymbol{\varepsilon}_g, \alpha_g) \quad \text{with} \quad Z_g = \mathbf{N}^\lambda_g \mathbf{L} + r\mathbf{N}^a_g \mathbf{A}. \tag{36}$$

As Π benefits from convexity, coercivity and semi-lower continuity properties, the differential inclusion (36) admits a unique solution for given values of $\boldsymbol{\varepsilon}_g$ and Z_g, denoted in an abstract format:

$$\alpha_g = \hat{\alpha}(\boldsymbol{\varepsilon}_g, Z_g). \tag{37}$$

For instance, consider again the case of the dissipation potential (18), i.e. $\Psi(\dot{a}) = k\dot{a} + I_{\mathbb{R}^+}(\dot{a})$. Then, according to (25)b, (17) and (19), the solution function $\hat{\alpha}$ is either equal to α^n_g or the unique root of a scalar equation:

$$\begin{cases} \text{if } -\dfrac{\partial \Phi}{\partial a}\left(\varepsilon_g, \alpha^n_g\right) + Z_g - r\alpha^n_g - k \leqslant 0 \quad \text{then } \alpha_g = \alpha^n_g, \\ \text{else find } \alpha_g \geqslant \alpha^n_g \quad \text{so that } -\dfrac{\partial \Phi}{\partial a}\left(\varepsilon_g, \alpha^n_g\right) + Z_g - r\alpha^n_g - k = 0. \end{cases} \tag{38}$$

Thus, the local problems (36) are very close to the integration of the purely local constitutive law since the only difference is the introduction of the affine term $Z_g - r\alpha_g$. Once α_g is computed, the stress tensor results from the definition (28). It reads in an abstract format:

$$\boldsymbol{\sigma}_g = \hat{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon}_g, Z_g) = \frac{\partial \Phi}{\partial \varepsilon}\left(\boldsymbol{\varepsilon}_g \hat{\alpha}(\boldsymbol{\varepsilon}_g, Z_g)\right). \tag{39}$$

3.2.3. Global balance equations

The spatial discretisation (33), the quadrature rule and the constitutive expressions (37) are then successively introduced into the Eqs. (28)–(30). First, the spatial discretisation (33) leads to:

$$\begin{cases} \boldsymbol{\delta}\mathbf{U}^T \displaystyle\int_\Omega \mathbf{B}^{uT} \boldsymbol{\sigma} = \boldsymbol{\delta}\mathbf{U}^T \mathbf{F}_{\mathbf{ext}} \quad \text{with} \quad \mathcal{W}'_{ext}(\mathbf{N}^u \boldsymbol{\delta}\mathbf{U}) = \boldsymbol{\delta}\mathbf{U}^T \mathbf{F}_{\mathbf{ext}}, \\ \boldsymbol{\delta}\mathbf{A}^T \displaystyle\int_\Omega c\mathbf{B}^{aT}\mathbf{B}^a\mathbf{A} + \boldsymbol{\delta}\mathbf{A}^T \int_\Omega \mathbf{N}^{aT}(\mathbf{N}^a\mathbf{A} - \alpha) + \boldsymbol{\delta}\mathbf{A}^T \int_\Omega c\mathbf{N}^{aT}\mathbf{N}^\lambda \mathbf{L} = 0, \\ \boldsymbol{\delta}\mathbf{L}^T \displaystyle\int_\Omega \mathbf{N}^{\lambda T}(\mathbf{N}^a A - \alpha) = 0. \end{cases} \tag{40}$$

The integrals are approximated through the quadrature rule (35), providing the corresponding expressions of the residual vectors $(\mathbf{R}^u, \mathbf{R}^a, \mathbf{R}^\lambda)$:

$$\begin{cases} \mathbf{R}^u = \displaystyle\sum_g w_g [\mathbf{B}^{uT}_g \boldsymbol{\sigma}_g], \\ \mathbf{R}^a = \displaystyle\sum_g w_g \left[c\mathbf{B}^{aT}_g \mathbf{B}^a_g \mathbf{A} + r\mathbf{N}^{aT}_g \left(\mathbf{N}^a_g \mathbf{A} - \alpha_{\mathbf{g}}\right) \mathbf{N}^{aT}_g \mathbf{N}^\lambda_g \mathbf{L}\right], \\ \mathbf{R}^\lambda = \displaystyle\sum_g w_g \left[\mathbf{N}^{\lambda T}_g \left(\mathbf{N}^a_g \mathbf{A} - \alpha_g\right)\right]. \end{cases} \tag{41}$$

The residual vectors $(\mathbf{R}^u, \mathbf{R}^a, \mathbf{R}^\lambda)$ are based on the stress tensor $\boldsymbol{\sigma}_g$ and the damage scalar α_g solutions of the pointwise constitutive relation as given in (37) and (39):

$$\begin{aligned} &\boldsymbol{\sigma}_g = \hat{\boldsymbol{\sigma}}\left(\mathbf{B}^u_g \mathbf{U}, \mathbf{N}^\lambda_g \mathbf{L} + r\mathbf{N}^a_g \mathbf{A}\right) \quad \text{and} \\ &\alpha_g = \hat{\alpha}\left(\mathbf{B}^u_g \mathbf{U}, \mathbf{N}^\lambda_g \mathbf{L} + r\mathbf{N}^a_g \mathbf{A}\right). \end{aligned} \tag{42}$$

Hence, $(\mathbf{R}^u, \mathbf{R}^a, \mathbf{R}^\lambda)$ are functions of the nodal unknowns $(\mathbf{U}, \mathbf{A}, \mathbf{L})$ only. Finally, application of any virtual vector $\delta\mathbf{U}$ (taking into account the

kinematical boundary conditions), $\delta\mathbf{A}$ and $\delta\ \mathbf{L}$ in (40) results in the following algebraic nonlinear system, the unknowns of which are the nodal vectors $(\mathbf{U},\mathbf{A},\mathbf{L})$:

$$\begin{cases} \mathbf{R}^u(\mathbf{U},\mathbf{A},\mathbf{L}) = \mathbf{F}_{\text{ext}}, \\ \mathbf{R}^a(\mathbf{U},\mathbf{A},\mathbf{L}) = \mathbf{0}, \\ \mathbf{R}^\lambda(\mathbf{U},\mathbf{A},\mathbf{L}) = \mathbf{0}. \end{cases} \tag{43}$$

The solution of the nonlinear system (43) is classically performed by means of Newton's method. It relies in particular on the consistent tangent matrix $\mathbf{K}$ which is derived in a straightforward way. It is symmetrical, thanks to the saddle-point formulation and it reads:

$$\mathbf{K} = \sum_g w_g \begin{bmatrix} \mathbf{B}_g^{u^T}\dfrac{\partial\hat{\boldsymbol{\sigma}}}{\partial\boldsymbol{\varepsilon}}\mathbf{B}_g^u & r\mathbf{B}_g^{u^T}\dfrac{\partial\hat{\boldsymbol{\sigma}}}{\partial Z}\mathbf{N}_g^a & \mathbf{B}_g^{u^T}\dfrac{\partial\hat{\boldsymbol{\sigma}}}{\partial Z}\mathbf{N}_g^\lambda \\ \text{sym.} & c\mathbf{B}_g^{a^T}\mathbf{B}_g^a + r\left(1 - r\dfrac{\partial\hat{\alpha}}{\partial Z}\right)\mathbf{N}_g^{a^T}\mathbf{N}_g^a & \left(1 - r\dfrac{\partial\hat{\alpha}}{\partial Z}\right)\mathbf{N}_g^{a^T}\mathbf{N}_g^\lambda \\ \text{sym.} & \text{sym.} & -\dfrac{\partial\hat{\alpha}}{\partial Z}\mathbf{N}_g^{\lambda^T}\mathbf{N}_g^\lambda \end{bmatrix}. \tag{44}$$

3.2.4. Shape functions and quadrature rule

Because of the straightforward treatment in (43) of the damage gradient term, the shape functions $\mathbf{N}^a$ should be continuous. The same holds for the displacement shape functions $\mathbf{N}^u$. By analogy with the results in [28] for poroelasticity, it is thought optimal to focus on displacement shape functions one order higher than damage shape functions; moreover, it avoids possible spurious stress oscillations [29]. This results in choosing polynomial Lagrange shape functions of degree 2 for the displacements and of degree 1 for the damage field.

The accuracy of the quadrature rule should be adapted to these sets of shape functions. Thus, the computation of $\mathbf{R}^u$ in (43) may rely on schemes with 8 Gauss points for hexahedra and 4 Hammer points for tetrahedra (4 Gauss points for quadrangles and 3 Hammer points for triangles). It is convenient to consider a single quadrature rule so as to avoid either projection from one set of integration points to another or too many solution of local problems. Since $\mathbf{N}^a$ consists of polynomials of degree one, the former set of integration points is satisfactory for the computation of $\mathbf{R}^a$ and $\mathbf{R}^\lambda$ as soon as $\mathbf{N}^\lambda$ is also piecewise polynomial of degree 1. Note that with such a set of integration points, the collocation discretisation of the field α in (34) corresponds to a spatial interpolation with shape functions that are also piecewise polynomials of degree one, denoted $\mathbf{N}^\alpha$.

Regarding the Lagrange multiplier field λ, a simple way to ensure the inf-sup condition consists in considering shape functions $\mathbf{N}^\lambda$ which are included in the set of shape functions $\mathbf{N}^\alpha$, i.e. $\mathit{Range}(\mathbf{N}^\lambda) \subset \mathit{Range}(\mathbf{N}^\alpha)$. Indeed, a lower bound can be obtained for the sup part by taking $a = 0$ and $\alpha = \lambda$:

$$\begin{aligned} &\inf_{\lambda\in\text{Range}(\mathbf{N}^\lambda)} \sup_{\substack{a\in\text{Range}(\mathbf{N}^a)\\ \alpha\in\text{Range}(\mathbf{N}^\alpha)}} \frac{(\lambda, a-\alpha)_0}{\|\lambda\|_0(\|a\|_1 + \|\alpha\|_0)} \\ &\geqslant \inf_{\lambda\in\text{Range}(\mathbf{N}^\lambda)} \sup_{\alpha\in\text{Range}(\mathbf{N}^\alpha)} \frac{(\lambda,\alpha)_0}{\|\lambda\|_0\|\alpha\|_0} \geqslant \inf_{\lambda\in\text{Range}(\mathbf{N}^\lambda)} \frac{(\lambda,\alpha)_0}{\|\lambda\|_0\|\lambda\|_0} = 1, \end{aligned} \tag{45}$$

where $(\cdot,\cdot)_0$ denotes the L^2 scalar product, $\|\cdot\|_0$ the corresponding L^2 norm and $\|\cdot\|_1$ the usual H^1 norm. In that way, the largest possible set for $\mathbf{N}^\lambda$ is then $\mathbf{N}^\alpha$.

In particular, the choice $\mathbf{N}^\lambda = \mathbf{N}^\alpha$ would result in exact satisfaction of the constraint $\mathbf{N}_g^a\mathbf{A} = \alpha_g$ for any point g. This corresponds to the choice made in [16]. Unfortunately, the corresponding consistent tangent matrix may not be invertible. Indeed, consider the case where the damage threshold is reached nowhere (elastic regime). Thanks to (38), the solution α_g to the local integration does not depend on $\boldsymbol{\varepsilon}_g$ nor Z_g. According to (44), it implies that the tangent matrix takes the simpler expression:

$$\mathbf{K}^{\text{EL}} = \begin{bmatrix} \mathbf{S} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \mathbf{C} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \quad \text{with} \quad \begin{cases} \mathbf{S} = \sum_g w_g \mathbf{B}_g^{u^T}\dfrac{\partial^2\boldsymbol{\Phi}}{\partial\boldsymbol{\varepsilon}^2}\mathbf{B}_g^u \\ \mathbf{M} = \sum_g w_g\left(c\mathbf{B}_g^{a^T}\mathbf{B}_g^a + r\mathbf{N}_g^{a^T}\mathbf{N}_g^a\right) \\ \mathbf{C} = \sum_g w_g\mathbf{N}_g^{a^T}\mathbf{N}_g^\lambda. \end{cases} \tag{46}$$

$\mathbf{S}$ is the elasticity matrix which is invertible (except for the points which are completely damaged). Therefore, $\mathbf{K}^{\text{EL}}$ is invertible if and only if the other 2×2 block matrix is invertible. As the matrix $\mathbf{M}$ is invertible, the kernel is defined as:

$$\begin{bmatrix} \mathbf{M} & \mathbf{C} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{A} \\ \mathbf{L} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Longleftrightarrow \begin{cases} \mathbf{A} = \mathbf{M}^{-1}\mathbf{C}\mathbf{L} \\ (\mathbf{C}^T\mathbf{M}^{-1}\mathbf{C})\mathbf{L} = \mathbf{0}. \end{cases} \tag{47}$$

In particular, it appears that $\mathbf{L}$ is left undefined if $n^\lambda \geqslant n^a$ (i.e. the number of degrees of freedom of $\mathbf{L}$ is larger than the one of $\mathbf{A}$); a condition that is unfortunately satisfied with $\mathbf{N}^\lambda = \mathbf{N}^\alpha$. We propose to choose $\mathbf{N}^\lambda = \mathbf{N}^a$: the Lagrange multipliers are also discretised by means of (continuous) Lagrange polynomials of degree 1. In that case, $\mathbf{C}$ becomes an invertible square matrix so that the tangent matrix $\mathbf{K}^{\text{EL}}$ is invertible too.

With such a discretisation for the Lagrange multiplier field, the constraint $\alpha = a$ is only weakly enforced, say "enforced in average". Because of the harmful characteristics of the energy density Π, a localisation phenomenon of α may still appear inside a finite element, though fulfilling the equality "in average" with the more regular field a. This may result in oscillations of α inside the elements as a consequence of the choice of shape functions. Fortunately, the penalty term in the collocated constitutive relation (36) controls these possible oscillations since it highly penalises any gap between α_g and $\mathbf{N}_g^a\mathbf{A}$. In that way, the dualisation takes care of the constraint "in average" while the penalty term controls the constraint at the finite element level.

An alternative approach to the discretisation choice $\mathbf{N}^\lambda = \mathbf{N}^a$ is proposed in [20]. It relies on a sole penalty method in order to enforce the constraint $\alpha = a$, instead of dualisation with Lagrange multipliers. In that case, the line and column corresponding to $\mathbf{L}$ in the matrix (47) vanish, so that $\mathbf{K}^{\text{EL}}$ is invertible. Moreover, such an approach is consistent since the limit case (for infinite penalty) corresponds to a dualisation with $\mathbf{N}^\lambda = \mathbf{N}^\alpha$ which fulfils the inf-sup condition. However, penalty methods have some drawbacks, among which the dependence of the solution on the penalty parameter (possibly problematic in the case of non convex problems) and the bad conditioning of the non linear system. Even though they can be overcome, dualisation generally results in more robust models from our experience in cohesive zone models and strain gradient damage laws. Here again, such an approach is preferred to a penalty method.

In summary, we make the following choice of shape functions: the displacement relies on P2-continuous polynomials, the damage field a and the Lagrange multipliers field λ on P1-continuous polynomials and the local damage field α on P1-discontinuous polynomials. The later corresponds to a collocation method on a set of 4 points (tetrahedron) or 8 points (hexahedron) which is respectively chosen as the set of Hammer or Gauss points so as to provide a quadrature rule. This discretisation is put into practice in Section 5 with satisfactory results.

4. Investigation of constitutive relations in an one-dimensional context

Despite the expression of gradient damage laws as a minimisation problem, computations encounter severe convergence

difficulties in practice. Thus, even for a simple brittle constitutive law of the Marigo type [30], the following facts are observed:

- The localisation bands narrow with increasing damage. It results in swapping from the loading regime to the unloading one since initially damaging points along the border of the band stop damaging while the band narrows. That may slow down the convergence. Moreover, because of these successive transitions from loading to unloading, time steps should be small enough in order to capture properly the transverse profile of the band, hence the amount of dissipated energy.
- Spurious snap-backs are observed [16], the size of which decreases with mesh refinement while their number increases (as observed with local laws). They preclude the application of an increasing load either because the computation simply does not converge or because the damage evolution exhibits some discontinuities with respect to the load, with consequences already mentioned in the previous point. Therefore a path-following technique is required but it needs very fine increments because of the large number of snap-backs. Hence an exploding computation time.

The combination of both effects, band narrowing and snap-back appearance, definitely precludes the application of the model to full-scale structures, in spite of the simplicity of the constitutive law (isotropy, elementary damage threshold, no crack closure, no irreversible strain). Our guess is that the undesirable properties not only result from the regularisation choice – the introduction of damage gradient – but also on the expression of the softening law itself. Therefore, the effects of variations of the softening part of the constitutive relations are analysed while preserving the peak stress value and the total amount of dissipated energy which are meaningful macroscopic quantities. The analysis is led in a 1D context for which a closed-form solution is available.

4.1. Constitutive laws

We restrict our attention to a bounded parameterisation of damage, a ranging from 0 (sound material) to 1 (ultimate damage). For the sake of simplicity, the dissipation potential is given by (18), i.e. $\Psi(\dot{a}) = k\dot{a} + I_{\mathbb{R}^+}(\dot{a})$, where the indicator function enforces irreversibility; it corresponds to the yield function (19), that is $f(y) = y - k$. Thus, the variety of the constitutive law lies in the free Helmholtz energy potential which may read for brittle materials:

$$\Phi(\boldsymbol{\varepsilon}, a) = \mathrm{A}(a)\mathrm{w}(\boldsymbol{\varepsilon}) + \omega(a) - ka \quad \text{with} \quad \mathrm{w}(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbf{E} : \boldsymbol{\varepsilon}, \tag{48}$$

where $w(\boldsymbol{\varepsilon})$ denotes the undamaged elastic energy. Note that the term $-ka$ is introduced into (48) without loss of generality since the function ω is let undefined at this stage; it enables some simplification of the expressions later on. As the energy is defined up to a constant, we set $\omega(0) = 0$ corresponding to zero free energy for the natural state. In particular, the stress and the damage driving force read:

$$\boldsymbol{\sigma} = \mathrm{A}(a)\mathbf{E} : \boldsymbol{\varepsilon}; \quad Y = -\mathrm{A}'(a)w(\boldsymbol{\varepsilon}) - \boldsymbol{\omega}'(a) + k. \tag{49}$$

For that kind of isotropic model, quadratic strain and stress measures are introduced:

$$w_\varepsilon = \frac{\boldsymbol{\varepsilon} : \mathbf{E} : \boldsymbol{\varepsilon}}{2}; \quad w_\sigma = \frac{\boldsymbol{\sigma} : \mathbf{E}^{-1} : \boldsymbol{\sigma}}{2} = A(a)^2 w_\varepsilon. \tag{50}$$

Note that in case of a homogeneous response (all fields are constant with respect to x), the strain and stress levels can be parameterised by the damage level as soon as the initial threshold is reached, thanks to the consistency condition $f(Y) = 0$:

$$w_e = -\frac{\omega'(a)}{A'(a)}; \quad w_\sigma = -\frac{\omega'(a)A(a)^2}{A'(a)}. \tag{51}$$

In particular, the initial damage threshold is defined as:

$$w_y = w_\varepsilon(a = 0) = w_\sigma(a = 0) = -\frac{\omega'(0)}{A'(0)}. \tag{52}$$

Actually, the generality of (48) is somewhat reduced because of the constraints on the functions A and ω:

(i) The stiffness A should be positive (elastic stability) and strictly decreasing (brittle damage) with $A(0) = 1$ (initial stiffness) and $A(1) = 0$ (complete loss of stiffness). Consequently, ω should be strictly increasing in order to preserve the positiveness of the (quadratic) strain measure w_ε, according to (51).

(ii) The class of generalised standard materials prescribes that the potential Φ be convex with respect to the damage a whatever the strain $\boldsymbol{\varepsilon}$. Therefore, both functions A and ω have to be convex. In particular, these conditions along with (i) ensures that in case of a homogeneous response, w_ε is increasing with damage which is a necessary (local) stability condition.

(iii) There exists an elastic response prior to damaging, that is the damage threshold w^y is not equal to 0, hence $\omega'(0) > 0$ and $A'(0)$ finite.

(iv) The condition of strain-softening requires that in case of a homogeneous response, w_σ be a decreasing function of the damage, as given by (51).

4.2. Closed-form solution for the 1D problem

We consider a bar of length $2L$ sufficiently large so that the boundary conditions do not interfere with the damage process zone. The load corresponds to a prescribed displacement at both ends of the bar. A symmetrical solution for the localisation band is expected; without loss of generality, the symmetry point is taken as the axis origin $x = 0$. Therefore, only half of the bar is studied; the domain of interest is $\Omega =]0L[$. Thanks to the equilibrium equation, the stress σ is homogeneous in Ω, corresponding to a homogeneous stress measure w_σ (50). The load level, i.e. the prescribed displacement, is parameterised in terms of the currently reached damage value at $x = 0$, say a_0, which is expected to be the maximum of the damage field at this stage. Such a parameterisation is valid as soon as the homogeneous preliminary elastic regime is overstepped, that is $w_\varepsilon \geqslant w_y$ somewhere. The damage field is then a function of the load parameter a_0 and the position x along the bar: $a(a_0, x)$.

Thanks to the expression of the dissipation potential Ψ, the value f of the yield function is given by (20). It reads:

$$f = -A'(a)w(\boldsymbol{\varepsilon}) - \omega'(a) + c\nabla^2 a. \tag{53}$$

The undamaged elastic energy can be expressed in terms of the (homogeneous) stress level according to (50), so that the yield function value is:

$$f = -\frac{A'(a)w_\sigma}{A(a)^2} - \omega'(a) + c\nabla^2 a. \tag{54}$$

The consistency condition (17) will then provide a differential equation which enables to find out the damage field. More precisely, we are looking for a damaged zone that develops in the area $[0b]$ where the boundary location b is a function of the load parameter a_0. This results in:

$$\begin{cases} f = 0 & \text{and} \quad \dot{a} \geqslant 0 \quad \forall\, x \in [0\ b], \\ f \leqslant 0 & \text{and} \quad \dot{a} = 0 \quad \forall\, x \in [b\ L]. \end{cases} \tag{55}$$

Now we make an assumption on the damage history that will be later on an essential condition for selecting a convenient law. The damage zone is assumed to spread, i.e. $b(a_0)$ is an increasing function, and remains an interval, which means that there is no damage unloading In that case, the differential system (55) can be integrated:

$$\begin{cases} f = 0 & \forall\, x \in [0\ b] \\ a = 0 & \forall\, x \in [b\ L] \end{cases} \quad \text{subjected to} \quad \begin{cases} \dot{a} \geqslant 0 & \forall\, x \in [0\ b], \\ f \leqslant 0 & \forall\, x \in [b\ L]. \end{cases} \tag{56}$$

Because of the interface conditions (15) and (16), the damage field is submitted to the following restrictions at the localisation band boundary:

$$a(a_0, b) = 0; \quad \frac{da}{dx}(a_0, b) = 0. \tag{57}$$

At $x = 0$, the load parameterisation and the symmetry condition read:

$$a(a_0, 0) = a_0; \quad \frac{da}{dx}(a_0, 0) = 0. \tag{58}$$

Finally, inside the localisation band $[0b]$, the damage field a obeys an ordinary differential equation of second order with respect to x subjected to four boundary conditions, that is two for the differential equation itself and two to set the boundary position b and the stress level w_σ. The differential equation $f = 0$ admits a first integral:

$$\frac{w_\sigma}{A(a)} - \omega(a) + \frac{c}{2}\left(\frac{da}{dx}\right)^2 = C. \tag{59}$$

The boundary conditions (57) provide the value of the integration constant $C = w_\sigma$. Moreover, the boundary conditions (58) prescribe the stress level:

$$w_\sigma = \frac{\omega(a_0)}{A(a_0)^{-1} - 1}. \tag{60}$$

The first order differential Eq. (59) can now be expressed as:

$$\frac{c}{2}\left(\frac{da}{dx}\right)^2 = \omega(a) - \omega(a_0)\frac{A(a)^{-1} - 1}{A(a_0)^{-1} - 1}. \tag{61}$$

The strict positiveness of the right-hand side of (61) can be demonstrated for $0 \leqslant a < a_0 < 1$. Indeed, it is equivalent to:

$$\forall\ a < a_0 \quad \frac{\omega(a)}{A(a)^{-1} - 1} > \frac{\omega(a_0)}{A(a_0)^{-1} - 1}. \tag{62}$$

This is ensured if the following function is strictly decreasing:

$$\begin{aligned} &[0\ a_0] \to \mathbb{R}, \\ &\quad a \mapsto \frac{\omega(a)}{A(a)^{-1} - 1}. \end{aligned} \tag{63}$$

The decreasing character is related to a strictly negative derivative which reads:

$$\forall\ a \in [0\ 1[\quad \frac{\omega'(a)A^2(a)}{A'(a)} > -\frac{\omega(a)}{A(a)^{-1} - 1}. \tag{64}$$

The former inequality is indeed satisfied, thanks to the properties (i) and (iv) of the local model. Therefore, coming back to (61), the differential equation can be expressed in a separated variable format:

$$\frac{da}{dx} = -\mathrm{G}(a_0, a)^{1/2} \quad \text{with} \quad \mathrm{G}(a_0, a) = \frac{2\omega(a_0)}{c}\left[\frac{\omega(a)}{\omega(a_0)} - \frac{\mathrm{A}(a)^{-1} - 1}{A(a_0)^{-1} - 1}\right]. \tag{65}$$

By taking into account the boundary condition (58), it admits the following implicit solution, with $x(a_0, a)$ the inverse function of $a(a_0, x)$ for a given load level a_0:

$$x(a_0, a) = \int_a^{a_0} \mathrm{G}(a_0, s)^{-1/2}\, ds. \tag{66}$$

In particular, thanks to the boundary condition (57), the boundary location is given by:

$$b(a_0) = x(a_0, 0) = \int_0^{a_0} \mathrm{G}(a_0, s)^{-1/2}\, ds. \tag{67}$$

Remember that the nonlinear differential system (56) is submitted to conditions. In the interval $[b\ L]$, the yield function is indeed negative since it has been proved that w_σ as given by (60) is a decreasing function of a_0, thanks to the decreasing character of (63). Regarding the positiveness of $\dot{a}$ in $[0\ b]$ it can be shown after straightforward but lengthy calculations that it is simply equivalent to the increase of $b(a_0)$ with a_0:

$$\frac{db}{da_0}(a_0) \geqslant 0 \Rightarrow \forall x \in [0b(a_0)]\ \dot{a}(a_0, x) = \frac{\partial a}{\partial a_0}(a_0, x)\dot{a}_0 \geqslant 0. \tag{68}$$

Therefore, the damage distribution given by (66) is valid for spreading damage zone (only), thus depending on the increasing character of (67).

4.3. Ultimate damage distribution

In general, the integral (66) does not admit an analytical expression so that it will be computed through a numerical integration. However, in the special case $a_0 = 1$, the stress level w_σ reaches 0, see Eq. (60), and according to (61), the differential equation reduces to:

$$\frac{c}{2}\left(\frac{da_u}{dx}\right)^2 = \omega(a_u) \tag{69}$$

with $a_u(x) = a(1, x)$ the corresponding ultimate damage distribution. Note that the stiffness function $A(a)$ is not involved in the ultimate damage distribution. Besides, the fracture energy G_f can be interpreted as the sum of the stored and the dissipated energy:

$$G_f = 2\int_0^L \left(\omega(a_u) + \frac{c}{2}\nabla a_u^2\right)dx = 4\int_0^L \omega(a_u)dx. \tag{70}$$

If we consider now a stored energy function of the form $\omega(a) = ka$ as will be used hereafter, the differential Eq. (69) admits the following solution:

$$\omega(a) = ka \to a_u(x) = \left(1 - \frac{x}{D}\right)^2 \quad \text{with} \quad D = \sqrt{\frac{2c}{k}}, \tag{71}$$

where D can be interpreted as half the width of the damaged zone, i.e. $b(1) = D$. In that case, the ultimate damage distribution is parabolic. It can be noticed that the damage field does no more respect the interface condition $[\![\nabla a \cdot \mathbf{n}]\!] = 0$ at $x = 0$, which corresponds to the "sharp profile" already observed in [16]. Finally, the knowledge of the fracture energy, the band width and the peak stress level w_y, which are macroscopic parameters, enables the identification of the internal parameters k, c and $A'(0)$. Indeed, application of (70), (71) and (52) leads to the following expressions:

$$\omega(a) = ka \to k = \frac{3}{4}\frac{G_f}{D}; \quad c = \frac{1}{2}kD^2; \quad A'(0) = -\frac{3}{4}\frac{G_f}{Dw_y}. \tag{72}$$

4.4. Selection of constitutive relations

Several families of constitutive relations have been analysed through the 1D case in order to sort out those which spread the

damage zone and do not exhibit snap-backs when focusing on the localisation band. Their properties will be highlighted through different plots based on the following definitions of normalised quantities:

$$\bar{\sigma}=\frac{\sigma}{\sigma_y};\bar{\varepsilon}=\frac{\varepsilon}{\varepsilon_y};\bar{x}=\frac{x}{d_c};\bar{b}=\frac{b}{d_c}\quad\text{with}\quad d_c=\sqrt{-\frac{A'(0)w_y}{2c}}. \tag{73}$$

Note that the characteristic length d_c has been chosen according to (72) in order to have $D = d_c$ in the case of the stored energy function $\omega(a) = ka$.

First, consider the following expressions, named from their authors:

$$\text{(LSB)}\quad A(a)=1-a\quad \omega(a)=k\left[a+\left(\frac{3}{2}+\gamma\right)^{-1}(1-a)\ln(1-a)\right], \tag{74}$$

$$\text{(BFM)}\quad A(a)=(1-a)^2\quad \omega(a)=k\left[\left(\frac{1}{3}+\gamma\right)a+\frac{1}{2}a^2\right], \tag{75}$$

$$\text{(LB)}\quad A(a)=1-a\quad \omega(a)=k\frac{a}{1+\gamma-a}. \tag{76}$$

The model LSB (74) is extracted from [18] and recast in the energy format. The BFM model (75) is derived from the regularised fracture formulation proposed in [21]. A linear term has been added in the energy function ω so as to introduce a damage threshold, absent in the original setting. Finally, the LB model (76) is the constitutive law presented in [16]. In each case, the softening parameter γ is submitted to $\gamma > 0$ in order to get strain-softening right after the elastic regime. As can be observed in the Eqs. (65)–(67), the parameters k (scaling of ω) and c have no influence on the variations of the damage band width, except in terms of scaling. Therefore, only the softening parameter γ has been sampled for the analysis of the three models LSB, BFM and LB. The results are plotted in Fig. 1a in terms of the evolution of the band width b versus the damage level a_0: none of these models lead to a widening band. Thus, they are prone to convergence difficulties as explained above.

The definition of a new family of constitutive laws has been undertaken on the basis of two ingredients: a power law for the stiffness function $A(a)$ since it seems to help somewhat in the BFM model and another parameterisation of damage $(1-a)\to(1-a)/(1+\gamma a)$ in the LB model which would correspond to an increasing value of c with damage growth if inserted in the minimisation problem (22). It results in the following power-law model:

$$\text{(LG)}\quad A(a)=\left(\frac{1-a}{1+\gamma a}\right)^m\quad \omega(a)=ka\quad 1\leqslant m\quad 0<\gamma. \tag{77}$$

It is shown in Fig. 1b that for $m = 1$, the band width does not grow. However, the objective is reached for $m = 2$ and $m = 3$ since in that cases, ranges of values for γ lead to an increasing band width, respectively around $\gamma \geqslant 2.8$ and $\gamma \geqslant 1.5$. Actually, the existence of a lower bound for γ corresponds to an upper bound on the band width D because of (72):

$$A'(0)=-m(1+\gamma)\Rightarrow D=\frac{3G_f}{4mw_y(1+\gamma)}. \tag{78}$$

If we consider characteristic values for concrete, see Table 1, the upper bound for D are not stringent since it corresponds for both cases to a band width $2D$ of about 130 mm. From now on, the value $2D = 100$ mm is chosen according to Table 1, valid for both values of m. The corresponding damage distribution are plotted in Fig. 1c as an illustration of the band width evolution.

Even though the band width is increasing, the damage rate distribution may concentrate around the centre of the band and may finally lead to mesh dependency. A way to measure this effect consists in evaluating the (normalised) density of energy remaining to (store and) dissipate at a given stage $\bar{e}(x, a_0)$ along with its integrated (normalised) contribution $\bar{E}(a_0)$ which measures the energy yet to dissipate:

$$\bar{e}(x,a_0)=\frac{1}{G_f}\left[\omega(a_u(x))+\frac{c}{2}\nabla a_u^2(x)\right]-\frac{1}{G_f}\left[\omega(a(x,a_0))+\frac{c}{2}\nabla a^2(x,a_0)\right],$$
$$\bar{E}(a_0)=2\int_0^L\bar{e}(x,a_0)dx. \tag{79}$$

We can then define a (normalised) average length of the active process $\bar{d}(a_0)$

$$\bar{d}(a_0)=\frac{2}{D\bar{E}(a_0)}\int_0^L e(x,a_0)x\,dx. \tag{80}$$

The plot in Fig. 1d shows that the less energy yet to dissipate, i.e. the closer to ultimate fracture, the smaller the average length of the active process. This phenomenon is less pronounced for $m = 3$ than for $m = 2$ and the possible final mesh dependency would therefore be postponed. However, one can hope that mesh dependency would happen sufficiently late so that the damage path is already well defined by the previous stage of damage, even for $m = 2$. This will be checked hereafter.

The last property to be analysed is the global response of the bar. Because of elastic unloading in the undamaged part of the bar, possible global snap-backs are expected. Nevertheless, it is interesting to see whether such snap-backs also appear at the scale of the damage band alone. Therefore, the following average strain is introduced:

$$\langle\bar{\varepsilon}\rangle(a_0)=\frac{1}{b(a_0)}\int_0^{b(a_0)}\bar{\varepsilon}(x,a_0)dx. \tag{81}$$

It can be observed on Fig. 1e that the average strain is an increasing function of the damage evolution for $m = 2$ and $m = 3$ on the contrary of what happens for instance with $m = 1.8$. Actually, a parametric study shows that the threshold corresponds to $m = 2$. For lower values of m, the response $(\langle\bar{\varepsilon}\rangle, \bar{\sigma})$ exhibits a snap-back. This is thought to explain the abovementioned micro snap-backs observed during damage propagation; it will be illustrated in Section 5.1. For $m \geqslant 2$, it appears that the smaller m, the more brittle the response; actually, it seems that ultimate fracture of the bar is reached with a finite average strain for $m = 2$ and only asymptotically for $m > 2$. This is also highlighted in Fig. 1f where it can be seen that a far larger opening of the damaged band is required with $m = 3$ than with $m = 2$ in order to dissipate most of the fracture energy. This may be a disadvantage for large values of m because the model could fail to describe complete fracture, resulting in process zones larger than expected. This feature will also be illustrated in Section 5.1.

In conclusion, the gradient damage power law (LG) seems to exhibit attractive (and complementary) features for $m = 2$ and $m = 3$. More realistic cases of damage propagation are presented in Section 5 so as to assess its effective properties.

5. Numerical validation

5.1. Numerical protocol

This section assesses the physical and numerical properties of the gradient damage formulation (27) based on the power-law constitutive relation (77). Three types of structures are studied:

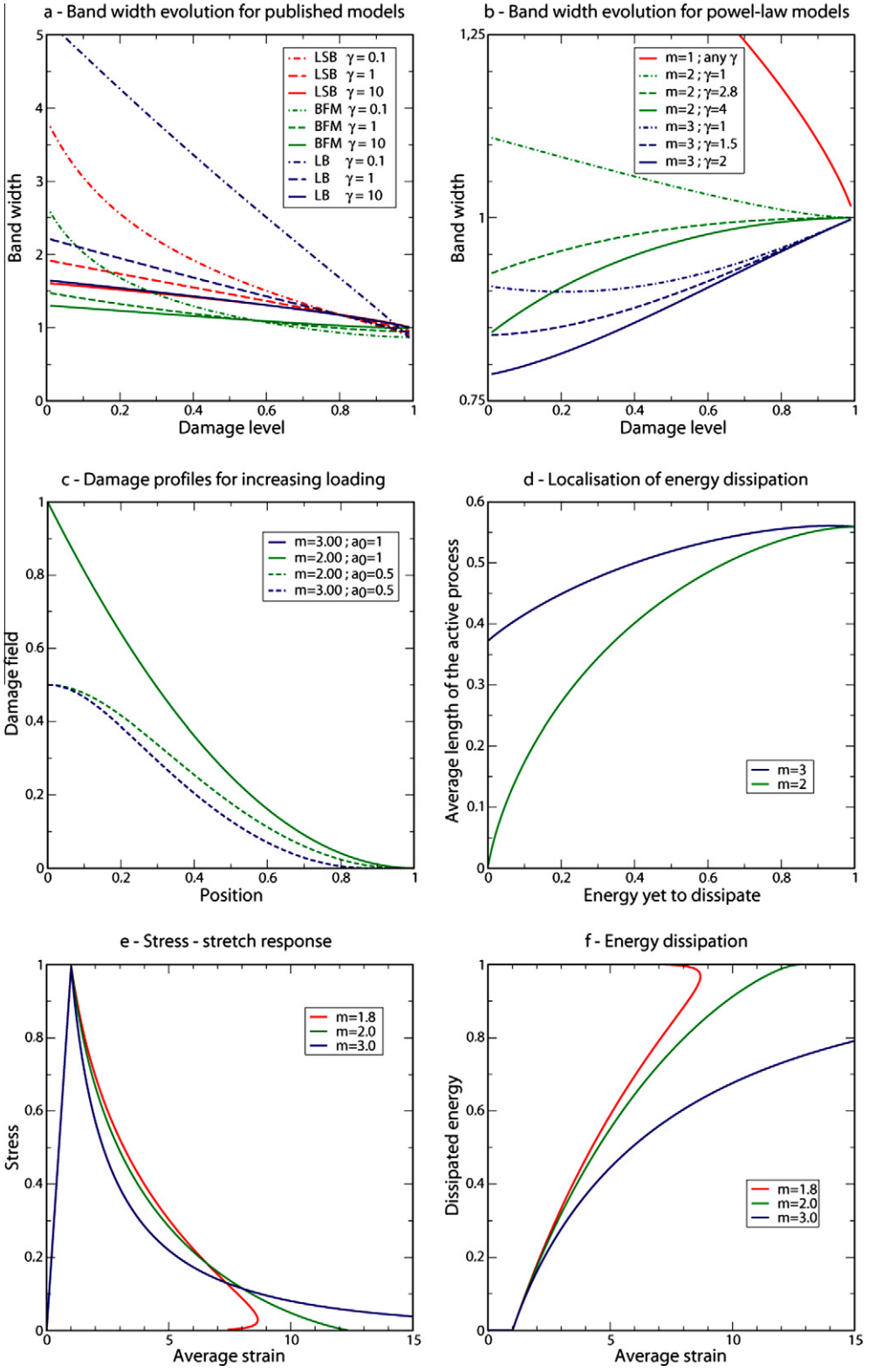

Fig. 1. Response of the 1D bar (normalised variables).

– a trapezium double cantilever beam (2D) to highlight the properties of (rectilinear and stable) crack propagation, including the sensitivity to mesh size;
– a perforated plate (2D) to estimate the robustness in presence of unstable multi-crack propagation and to quantify the sensitivity to mesh orientation;
– an L-shaped panel (3D) to give a first insight on crack path prediction for 3D loading.

Considering civil engineering as a possible target for such a formulation, the characteristic size of the previous structures reaches several meters. It is then reasonable to consider an internal

Table 1
Material parameters representative of concrete.

Observable quantities	Young modulus	$E = 30000$ MPa
	Poisson ratio	$\nu = 0.2$
	Peak stress	$\sigma_y = 3$ MPa $\rightarrow$ $w_y = 1.5 \times 10^{-4}$ MPa
	Fracture energy	$G_f = 0.1$ N/mm
	Band width	$2D = 100$ mm
Power-law identification:	Dissipation potential slope	$k = 1.5 \times 10^{-3}$ MPa
	Softening parameter	$m = 2 \rightarrow \gamma = 4$; $m = 3 \rightarrow \gamma = 2.33$
	Regularisation weight	$c = 1.875$ N

length $2D = 100$ mm, small compared to the characteristic structural length. The other material parameters are also representative of concrete and given in Table 1. As the model does not account for the dissymmetry between tension and compression nor possible crack closure, the considered loading are restricted to tension and shear modes, in order to focus on the properties of regularisation rather than on the deficiency of the simple constitutive law. We refer to [31] for more complex laws that may be cast in the present framework while taking into account microscopic crack closure and anisotropy. In addition to the material parameters, the augmented Lagrangian formulation involves a penalty parameter. The chosen value is $r = 10^3 k$, with k the dissipation potential slope, representative of the dissipated energy. It appears sufficient to control the possible oscillations inside the finite elements while not slowing down Newton's method through an excessive curvature.

Because of the brittle character of the material behaviour, instabilities are expected. Therefore, a path-following method should be applied to control the load magnitude in order to follow the possible snap-backs resulting from these material instabilities [32,33]. The path-following function is described in [34]; the choice relies on its good properties for local damage models and cohesive crack propagations. Its main idea consists in controlling the maximal damage increase $a^{n+1} - a^n$, which enforces the selection of a dissipative branch (propagation of damage). In practice, the maximal damage increment is set to 0.1, which means that for each load increment, the damage increment reaches 0.1 somewhere in the structure and does not exceed 0.1 anywhere else. Because Newton's method is not unconditionally convergent, a reduction of this load parameter is allowed. A simple strategy is applied: whenever Newton's method does not converge after 15 iterations, the load increment is (recursively) divided into 4 sub-increments.

The last numerical parameters concerns the convergence criterion, that is the accuracy of the solution of the nonlinear algebraic system (43). The criterion provides a nodal reference for each of the three families of equations, in comparison to which each component of the residual should be small, set to a factor $\varepsilon = 10^{-3}$ in the following simulations. If n denotes the dimension of the problem (2 for 2D, 3 for 3D) and h denotes the size of an element, then the nodal references are $h^{n-1}\sigma_y$ for $\mathbf{R}^u - \mathbf{F}_{ext}$, $h^n k$ for $\mathbf{R}^a$ and h^n for $\mathbf{R}^{\dot{z}}$, where the peak stress σ_y and the dissipation potential slope k are given in Table 1. Though this is not an essential feature of the model, the convergence criterion has been summed up here for the sake of reproductability of the numerical experiments. Any other convergence criterion that distinguishes the three families of components could have been applied.

Note that the computations described hereafter have been performed with Code_Aster, an open-source finite element software [35], which also accounts for reproductability.

5.2. Stable rectilinear crack propagation

The first numerical experiment investigates the properties related to the propagation of a damage zone. The structure and the loading have been designed in order to trigger the stable propagation of a rectilinear damage band. It consists of a pre-cracked trapezium double cantilever beam submitted to opening prescribed displacements as depicted in Fig. 2. A trapezium shape has been preferred to the classical straight shape of a DCB so as to avoid bifurcation from the rectilinear path.

Qualitatively, the numerical simulation leads indeed to the rectilinear propagation of a damage band which initiates at the crack end, see Fig. 3a. Moreover, the band width is about 100 mm, in agreement with the prediction of the 1D problem in Section 4. On Fig. 3e, the closeness with the 1D prediction is confirmed by the damage profile along a cross-section (dashed line on Fig. 3a). In particular, it spreads over several finite elements as expected from the regularisation properties. Besides, the fact that the band width remains constant along the damage zone should be pointed out as an interesting feature of gradient damage models since this property is not satisfied without additional complexity when regularising the strain field, see [17] for instance.

Quantitatively, the experiment has been led with three values of the power-law exponent m. The results are plotted on Fig. 3b

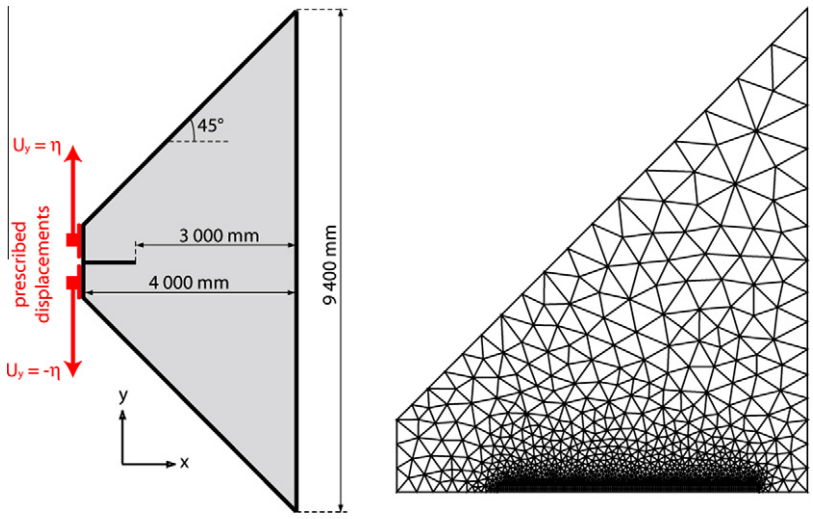

Fig. 2. Trapezium structure.

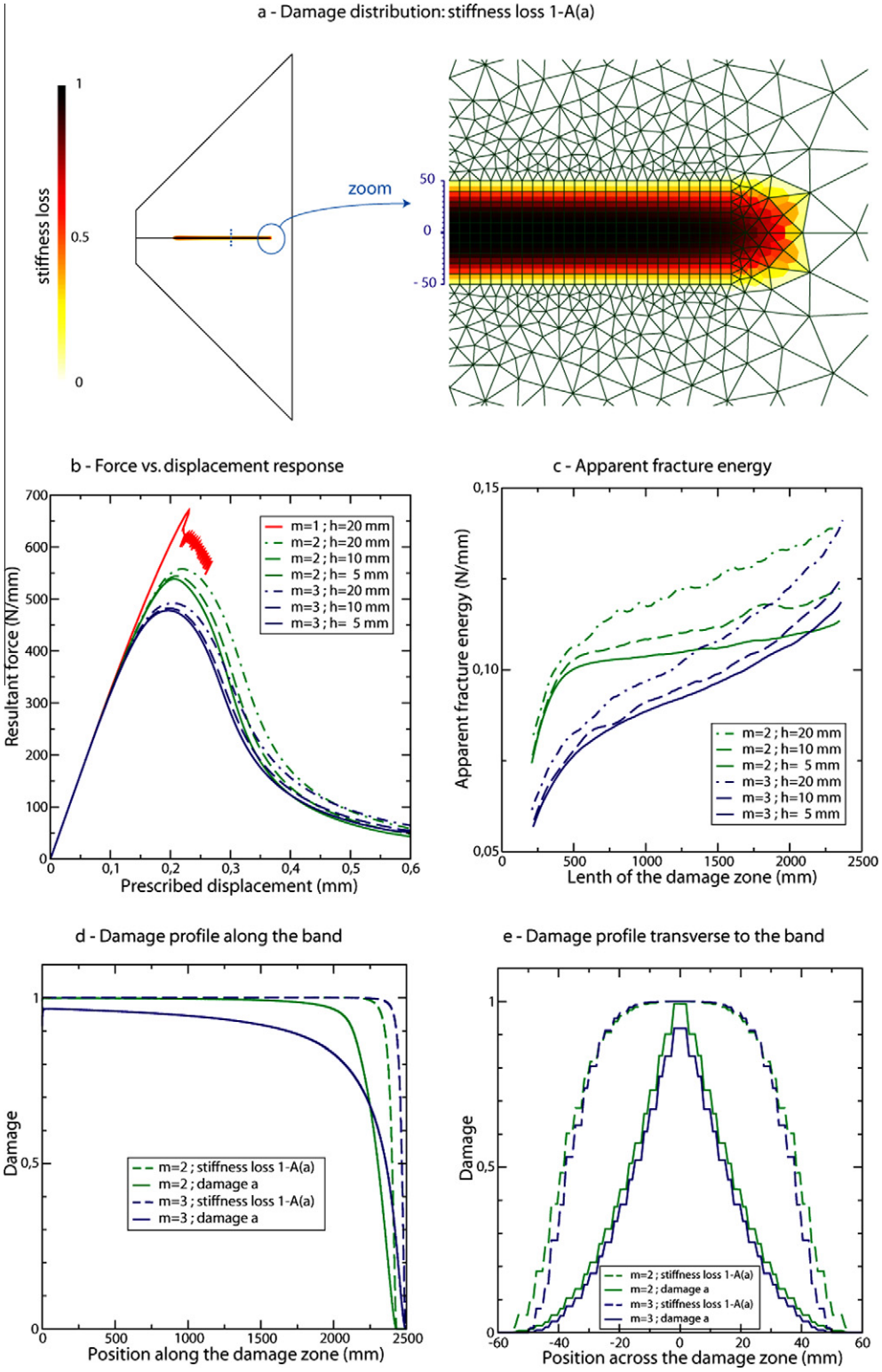

Fig. 3. Response of the trapezium structure.

in terms of force vs. displacement. For $m = 1$, it appears that the response exhibits many small snap-backs and ends up with lack of convergence far before any ultimate load. The results observed with classical models are retrieved. However, the guess of Section 4 is confirmed: with an exponent greater or equal to 2 ($m = 2$ and $m = 3$), the response is smooth and the computation can be led without difficulty up to ultimate failure (here, it has been stopped when compressive damage occurs at the right end of the specimen). The sensitivity to mesh size has also been analysed on the basis of three mesh refinements: 5, 10 and 20 elements under the band width respectively corresponding to $h = 20$ mm, $h = 10$ mm and $h = 5$ mm. Of course, a dependence is observed, slightly reduced in the case of $m = 3$ compared to $m = 2$. However, the results seem to converge and the discrepancy in terms of peak force between the coarse and the fine meshes is less than 4%, which is quite acceptable. In particular, the coarse mesh appears

to be a reasonable compromise between accuracy and efficiency for 3D simulations.

The 1D analysis concluded that higher exponents of the power law could reduce mesh dependency but may fail to describe the ultimate phase of fracture because of "excessive ductility". This is confirmed by the energetic analysis of the trapezium specimen. More precisely, on the one hand, an equivalent crack length is measured from the damage distribution, where the pseudo-crack end is arbitrarily set to a stiffness loss $A(a) = 0.5$. One the other hand, the amount of dissipated energy is computed for each load increment. This enables the computation of the rate of dissipated energy with respect to the crack length extension: in Griffith's theory of fracture, this corresponds exactly to the fracture energy. The results are plotted on Fig. 3c. Again, the results seem to converge with mesh refinement. Moreover, for the exponent $m = 2$, the damage process reaches a relatively stationary state with a rate of dissipated energy close to $G_f = 0.1$ N/mm. This is a very interesting feature which places the damage gradient theory as a compatible extension of Griffith's theory (used in most industrial applications). On the contrary, the exponent $m = 3$ does not lead to a stationary state for the considered propagation length of 2500 mm.

The preference for the lower exponent $m = 2$ is confirmed by the quantitative analysis of the damage distribution along the crack path, see Fig. 3d. Even though the stiffness loss exhibits a sharp front, the distribution of damage is smoother. However, the extension of the process zone (defined as $0 < a \leqslant 0.95$ for instance) remains far lower with $m = 2$ (about 375 mm) than for $m = 3$ (over than 2 m). Compared to the characteristic length of the target structures, this hints for the choice $m = 2$. In that case, it should be noticed that the size of the process zone is compatible with the expression provided in [36] for a cohesive zone model:

$$L_{CZM} \approx \frac{9\pi}{32} \frac{E}{(1-\nu^2)} \frac{G_f}{\sigma_y^2} \rightarrow \text{N.A.} \quad L_{CZM} \approx 295 \text{ mm}. \tag{82}$$

Therefore, thanks to the desirable properties of compatibility with Griffith's theory and cohesive zone models, the forthcoming numerical experiments are established with a power-law the exponent of which is $m = 2$.

5.3. Complex propagations

For the scale $2D = 100$ mm which is considered here, the damage power-law with exponent $m = 2$ exhibits advantageous properties in terms of damage distribution and fracture energy. However, the analysis was based (i) on a stable propagation of the damage zone and (ii) on a mesh the orientation of which was adapted to the crack path. Therefore, this sections aims at:

- demonstrating the robustness of the approach in case of unstable multi-crack propagation;
- estimating the dependence on mesh orientation, which complements the previous analysis of convergence with the mesh size.

The latter point appears necessary since it has been pointed out in Section 4.4 that the power-law model with exponent $m = 2$ tends to ultimately concentrate the active damage zone, see again Fig. 1d, resulting in possible mesh dependency.

The numerical evaluation is led on a perforated plate subjected to tension on its opposite faces, see Fig. 4. As the hole is off centre, the damage zone is expected to propagate on the shortest side (right side), then on the other one (left side), i.e. multi-crack propagation. Moreover, computations with a cohesive zone model [34] have exhibited a propagation of unstable nature. Thus, both difficulties are triggered in the perforated plate case, multi-cracking and instabilities.

The simulation is first performed with the unstructured mesh depicted in Fig. 4, which is refined in the area where the cracks are expected with minimal mesh size $h = 20$ mm. The computation is stopped when compressive damage occurs (on the left wall). It appears that the results are consistent with the CZM computations. Two damage zones propagate, the first one across the short side (right) then the other across the large side (left), see Fig. 5a. Moreover, both are unstable, which corresponds to the snap-backs on the force–displacement response plotted in Fig. 5b. It should also be noticed that no transverse enlargement of the damage zone is observed in spite of the opening of the right ligament after its complete failure. As mentioned previously, this is worth pointing out compared to strain regularisation techniques [17]. Finally, these results prove that combining the present gradient damage formulation with the path-following technique [34] enables computation of complex damage zone propagations.

In order to study the sensitivity to mesh orientation, a focus is put on the right side damage zone. Several meshes are applied in that area, either unstructured (triangles) or structured along a privileged direction with angle 20° and 40° with respect to a horizontal crack path (quadrangles). Besides, two mesh sizes are considered in the oriented area: $h = 10$ mm and $h = 20$ mm. The results in

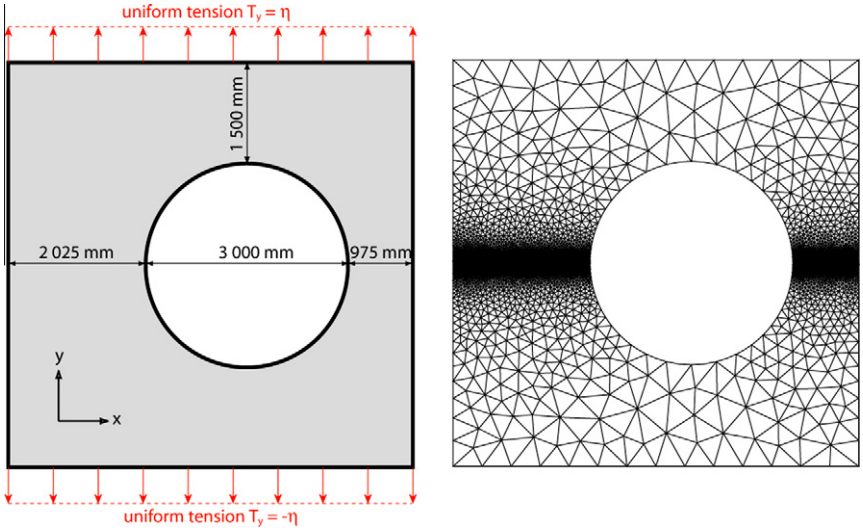

Fig. 4. Perforated plate.

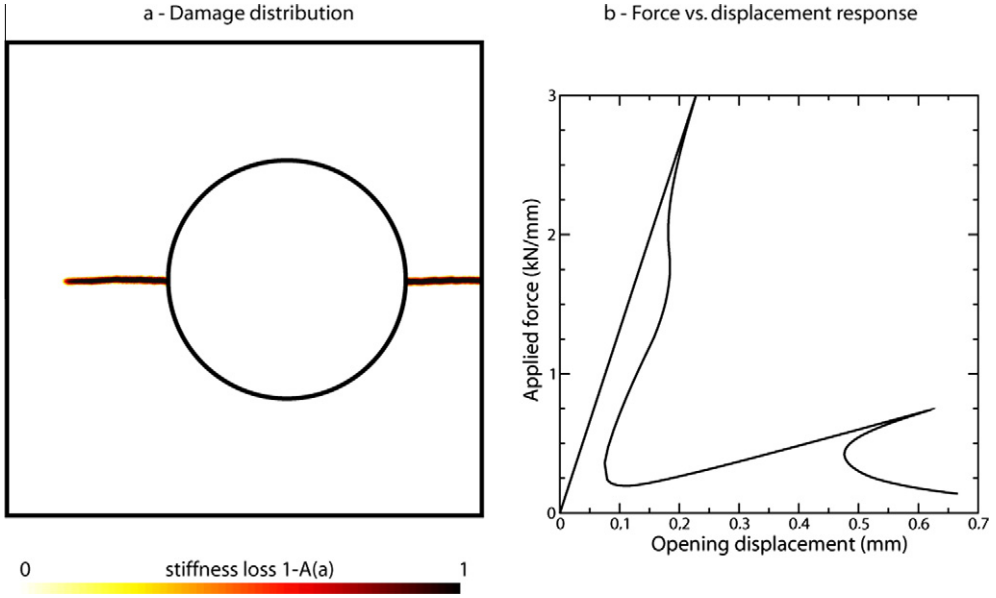

Fig. 5. Response of the perforated plate.

terms of damage (stiffness loss) distribution are presented in Fig. 6. for the six cases. A slight deviation from a horizontal crack path is observed, whatever the mesh. But along the one meter long ligament, it does not exceed 2° with the coarse mesh and 1° with the fine one. The convergence with mesh refinement is obvious. Moreover, the coarse mesh appears again as a good trade-off between accuracy and efficiency for 3D simulations.

5.4. 3D crack path prediction

The previous 2D simulations have shown that the damage power-law with exponent $m = 2$ is satisfactory regarding the physical predictions as well as the lack of spurious mesh dependency. Now, a 3D simulation is performed in order to check the capability of the gradient formulation in case of full-scale computations.

The considered structure is an L-shaped panel depicted in Fig. 7, inspired by a benchmark led in [37]. Clamped at one end, it is submitted to shear and opening displacements at the other one. According to 2D results, a crack is expected to propagate from the corner (even though the load is obviously not the same in 2D). Therefore, the mesh in the corresponding area is more refined than far from the corner, see Fig. 7 again: the finest elements are 25 mm large (4 elements under the expected band width) while the coarsest are 100 mm large. The refinement is based on the acceptable results obtained previously with 20 mm large elements. Such a discretisation results in about 370,000 nodal degrees of freedom (including displacements, nodal damage and Lagrange multipliers). The computation is stopped when compressive damage occurs on the left wall of the structure.

Qualitatively, the good properties observed in 2D still prevail. The global response in term of forces vs. prescribed displacement is plotted on Fig. 8b: it is very smooth; in particular, no spurious snap-back is observed. Besides, Fig. 8a shows that the damage (stiffness loss) is concentrated in a localisation band the width of which still remains close to 100 mm, as expected. The initiation of compressive damage is also apparent on the left wall. Finally, Fig. 8c to Fig. 8f illustrate in several ways the rotation of the band from horizontal (along the corner) to a slant configuration resulting from shear load and the curved damage path due to the opening component of the loading (already observed in 2D).

This demonstrates the capability of the model to predict complex fracture modes with a mesh refinement that remains compatible with practical studies.

5.5. Numerical convergence

Up to now, the analysis has been focused on the results and their numerical stability with respect to mesh refinement. It is also of practical interest to check the convergence of Newton's algorithm.

First, it appears that the initiation of the damage zone generally requires a finer damage increment (path-following method) than the propagation regime. Indeed, several subdivisions of the standard damage increment of 0.1 are necessary, resulting in a damage increment of about 0.0015. However, once the initial damage zone is set, the standard damage increment of 0.1 is sufficient; it corresponds to a propagation of the damage zone of 2–3 cm per increment, as observed for the three considered structures.

In this propagation regime, Newton's method requires a limited number of iterations in order to reach the convergence threshold, as presented in Fig. 9 for the exponent $m = 2$. The number of iterations does not depend on the problem size nor its 2D or 3D character. Thus, with the coarse mesh, 5–6 iterations per load increment are necessary in average. This is quite reasonable, close to what happens with classical nonlinear laws such as hardening plasticity. It appears also that the number of iterations increases with mesh refinement. Nevertheless, with a very fine mesh (20 elements under the band width), the average number of iterations remains less than 10, which is still acceptable in our opinion (in addition to the fact that such refined meshes do not seem necessary regarding accuracy).

This demonstrates that Newton's method combined with a path-following technique is well suited to the solution of the non linear system resulting from the gradient formulation of the power-law damage model.

6. Experimental validation

It has been pointed out above that the damage nonlocal model is compatible with coarser analyses based on Griffith's theory or cohesive zone models. Therefore, it benefits from the large basis

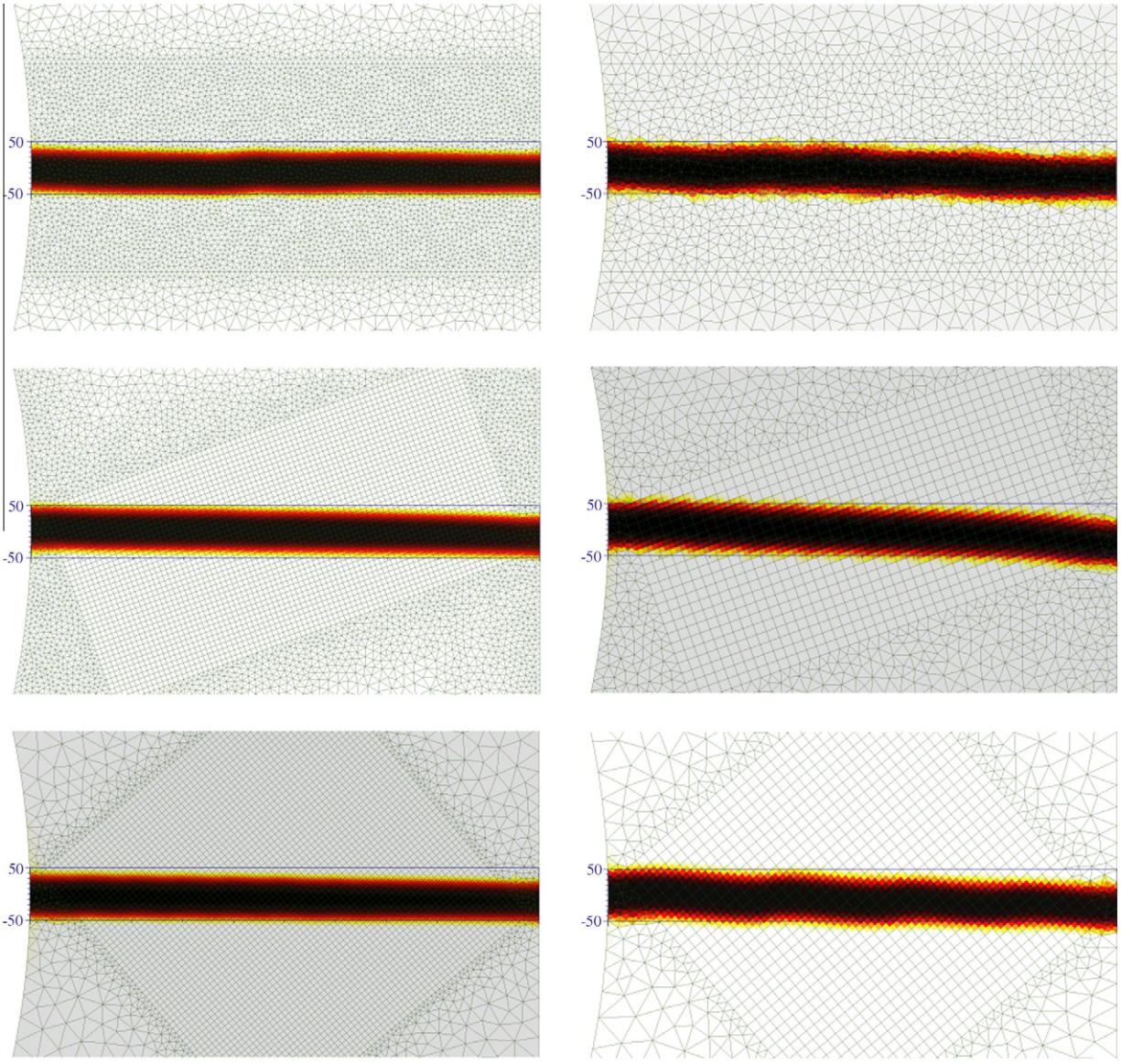

Fig. 6. Sensitivity to mesh orientation.

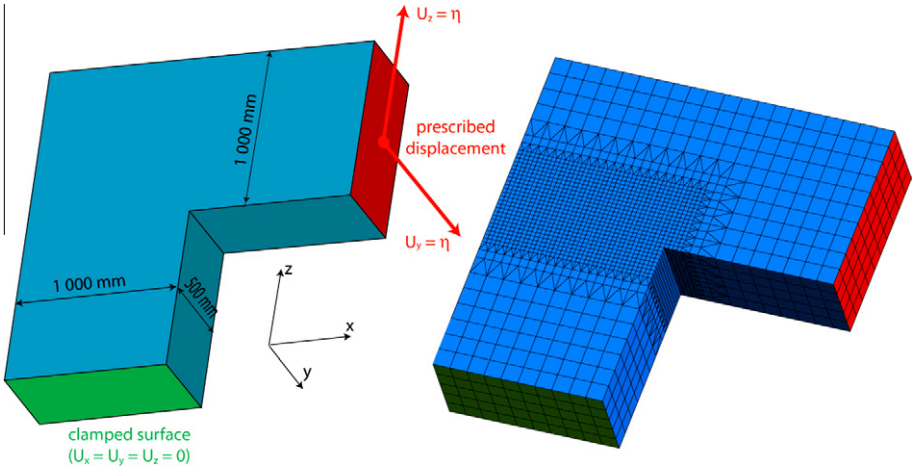

Fig. 7. L-shaped panel.

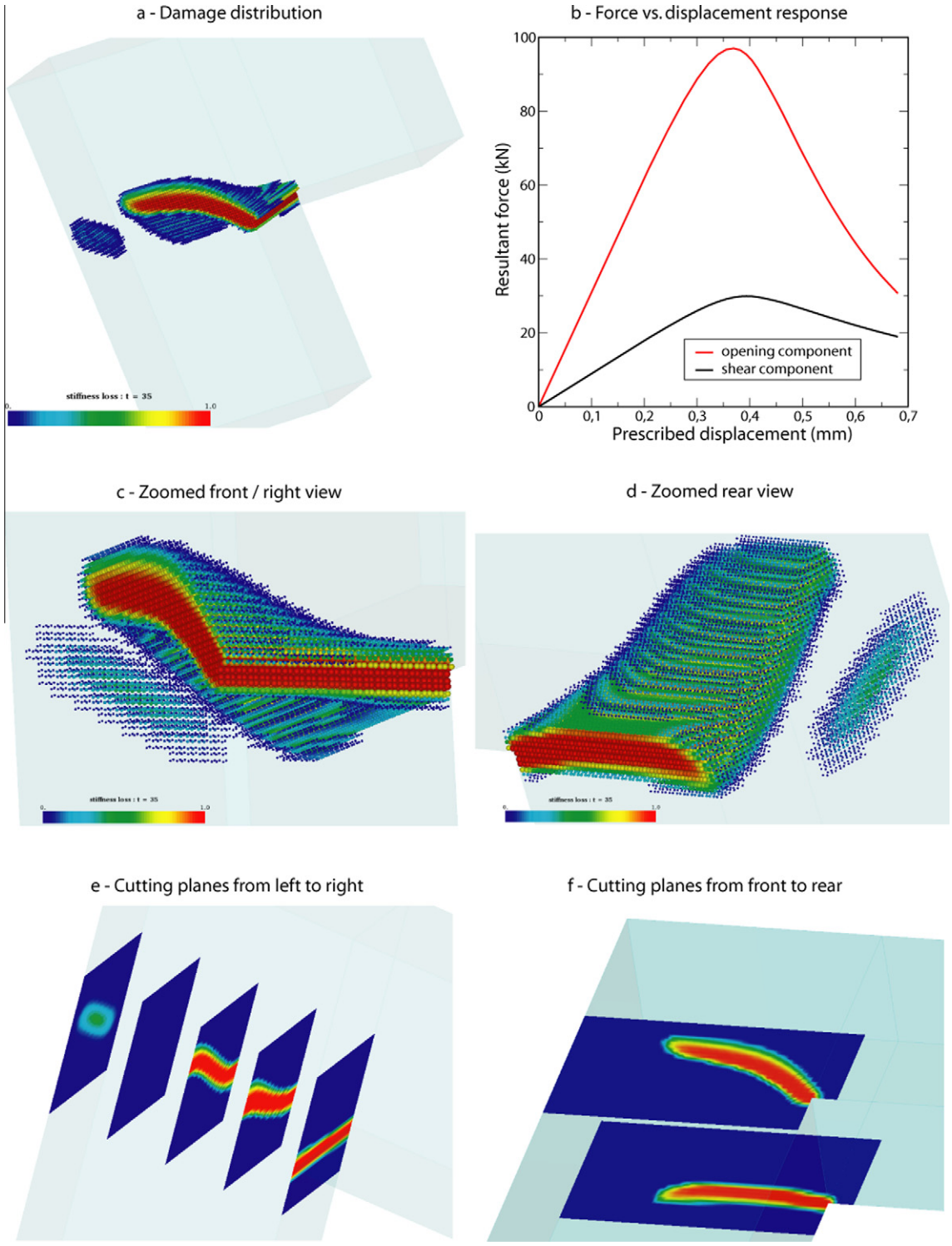

Fig. 8. Fracture of the L-shaped panel.

of validation of these theories, even though additional work has to be done so as to confirm and calibrate the effective relation between the approaches.

Moreover, comparisons with experimental results are complementary to assess the physical validity of the proposed nonlocal law. The choice of reference specimens should then cope with (i) the aim of describing large structures, (ii) a focus on the constitutive behaviour itself without any other modelling complexities such as reinforcement and (iii) the limited domain of validity of the simple brittle law (unphysical compressive damage, no crack closure, isotropy). Remember that the latter has been adopted in order to separate the numerical difficulties related to gradient nonlocality and the modelling issues of brittle material constitutive behaviour. Taking into account the previous requirements, the concrete crack-line wedge-loaded double cantilever beam (CLWL-DCB) with diagonal compression loading appears to be a convenient specimen in order to get a first estimation of the validity of the proposed formulation.

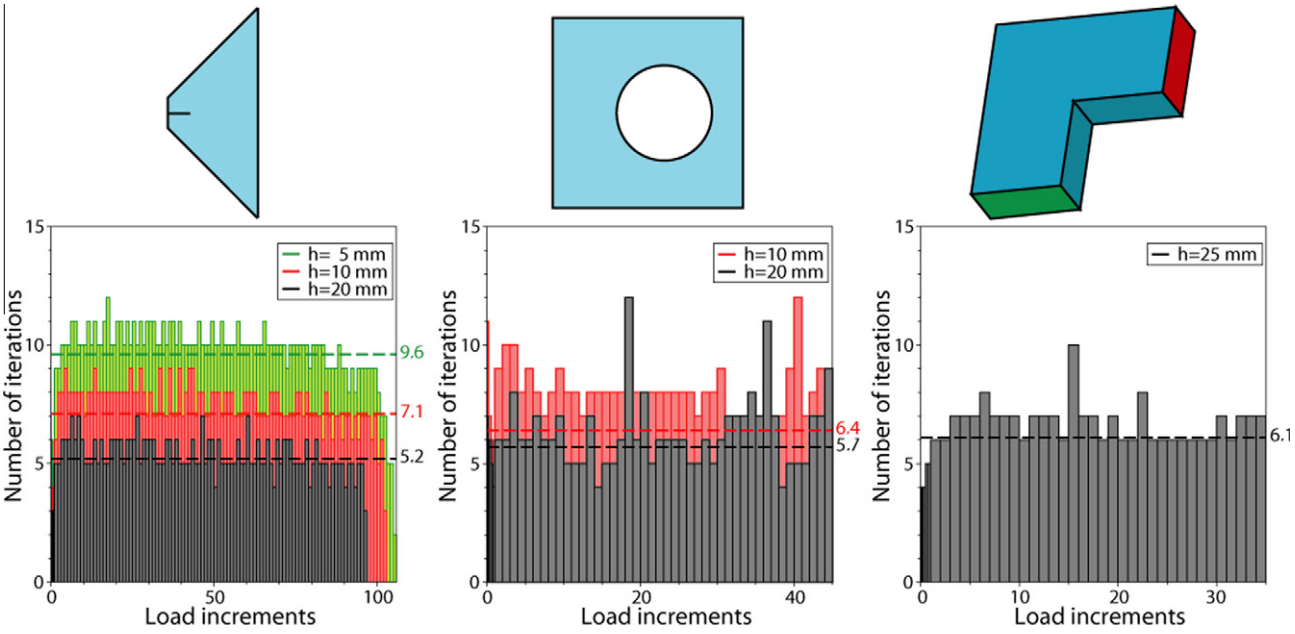

Fig. 9. Convergence of Newton's method.

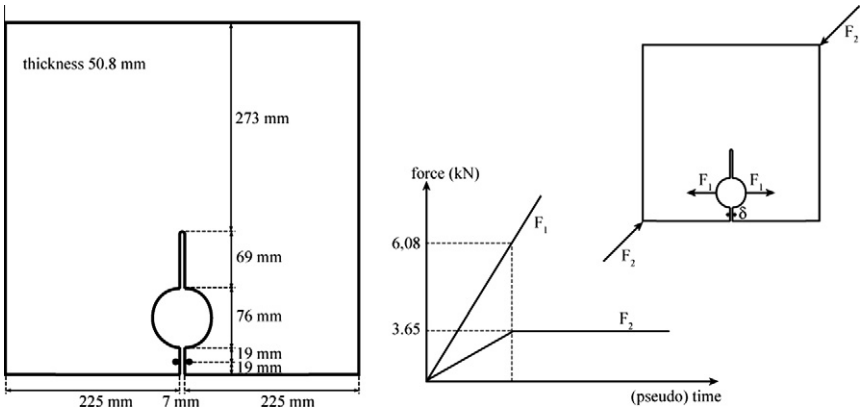

Fig. 10. Crack-line wedge-loaded double cantilever beam.

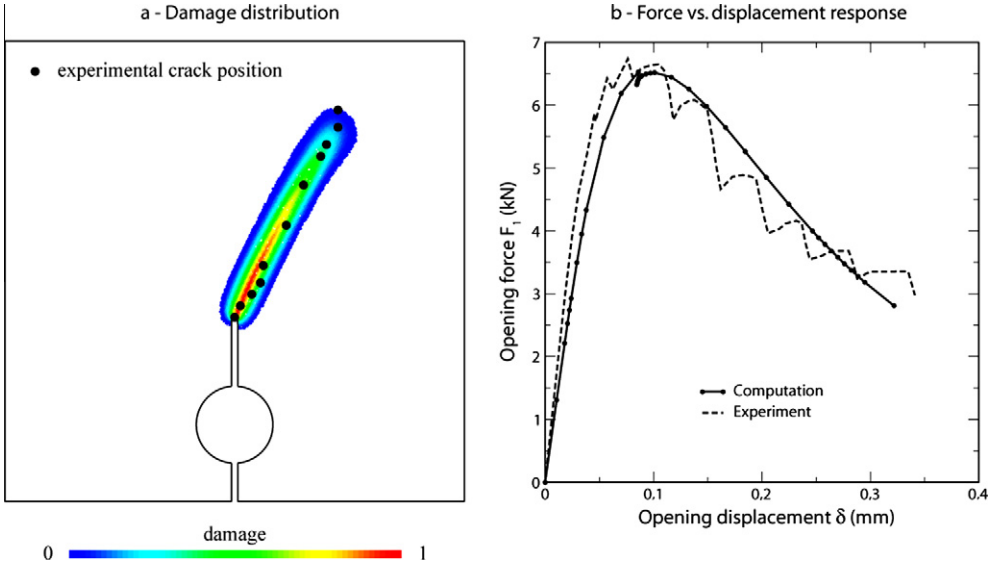

Fig. 11. Response of the CLWL-DCB.

A detailed description of the CLWL-DCB specimen and corresponding experimental results are provided in [38]. They are summed up in Fig. 10. No identification has been performed: the material parameters are the same as in the previous section with a power law exponent $m = 2$, see Table 1, except for the characteristic length which is chosen twice smaller $D = 50$ mm because of the smaller size of the specimen (in order to avoid interactions between the damage pattern and the structure). Thanks to the coarse comparison which is aimed at, a plane strain assumption is made. The computational strategy presented in Section 5 is applied again. The results are presented in Fig. 11 in terms of crack path prediction and crack opening displacement vs. applied force. Although no tuning of the material parameters has been performed, a reasonable agreement with the experimental measures is observed which confirms the qualitative and quantitative capabilities of the model.

7. Summary

Damage nonlocal constitutive laws involving the gradient of the damage field have been cast into a variational formulation. The introduction of an augmented Lagrangian has enabled a numerical treatment which:

- relies on a mixed finite element the unknowns of which are the displacements, the damage and Lagrange multipliers;
- condenses the high nonlinearities in pointwise nonlinear equations that correspond to the integration of the local constitutive relation perturbed by an additional affine term.

In that way, the impact of the formulation has been reduced to the design of a new finite element and a slight modification of the integration procedure of the constitutive law. In particular, classical solution algorithms such as Newton's method and path-following techniques are still available without any alteration.

Despite this numerical setting, the convergence of computations may be dramatically slow. This has been related to the fact that localisation bands narrow with increasing damage and to the appearance of micro snap-backs. On the basis of a 1D closed-form solution, it has been showed that these drawbacks can be overcome. In particular, a class of brittle damage laws has been proposed which removes them while preserving essential features: damage threshold, strain-softening, irreversibility, zero residual stiffness. Moreover, it relies on macroscopic material parameters only: the critical stress, the fracture energy and the band width.

Several numerical simulations have shown the characteristics of the model. It is compatible with former predictions based on Griffith's theory or cohesive cracks, which is of paramount industrial importance. It does not suffer from spurious mesh-dependency, neither regarding the mesh size nor the mesh orientation. It is able to cope with complex situations such as unstable multi-crack problems. It is sufficiently robust and efficient so as to perform full-scale 3D simulations, thanks to good convergence properties of Newton's algorithm. A comparison with experimental results finally confirms the physical validity of the approach.

References

[1] J. Lemaitre, J.-L. Chaboche, Mécanique des Matériaux Solides, Dunod, Paris, 1988.

[2] A. Benallal, R. Billardon, G. Geymonat, Bifurcation and localization in rate-independent materials. Some general considerations, in: Q.S. Nguyen (Ed.), Bifurcation and stability of dissipative systems, CISM Courses and Lectures, vol. 327, Springer Verlag, 1993, pp. 1–44.

[3] Z.P. Bazant, T.B. Belytschko, T.P. Chang, Continuum theory for strain-softening, J. Engng. Mech. Div. ASCE 110 (1984) 1666–1692.

[4] W.J. Drugan, J.R. Willis, A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites, J. Mech. Phys. Solids 44 (1996) 497–524.

[5] S. Andrieux, M. Joussemet, E. Lorentz, A class of constitutive relations with internal variable derivatives: derivation from homogenization, C.R. Acad. Sci. Paris 323 série IIb (1996) 629–636.

[6] E. Lorentz, S. Andrieux, Analysis of nonlocal models through energetic formulations, Int. J. Sol. Struct. 40 (2003) 2905–2936.

[7] G. Pijaudier-Cabot, Z.P. Bazant, Nonlocal damage theory, J. Engrg. Mech., ASCE 113 (1987) 1512–1533.

[8] R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans, J.H.P. de Vree, Gradient-enhanced damage for quasi-brittle materials, Int. J. Num. Methods Engrg. 39 (1996) 3391–3403.

[9] N. Triantafyllidis, E.C. Aifantis, A gradient approach to localization of deformation. I –Hyperelastic materials, J. Elast. 16 (1986) 225–237.

[10] R. Chambon, D. Caillerie, T. Matsuchima, Plastic continuum with microstructure, local second gradient theories for geomaterials: Localization studies, Int. J. Solids Struct. 38 (2001) 8503–8527.

[11] H.B. Mühlhaus, E.C. Aifantis, A variational principle for gradient plasticity, Int. J. Solids Struct. 28 (1991) 845–857.

[12] T. Svedberg, K. Runesson, A thermodynamically consistent theory of gradient-regularized plasticity coupled to damage, Int. J. Plast. 13 (1997) 669–696.

[13] M. Frémond, B. Nedjar, Damage, gradient of damage and principle of virtual power, Int. J. Solids Struct. 33 (1996) 1083–1103.

[14] G. Pijaudier-Cabot, N. Burlion, Damage and localisation in elastic materials with voids, Mech. Coh. Frict. Mat. 1 (1996) 129–144.

[15] E. Lorentz, S. Andrieux, A variational formulation for nonlocal damage models, Int. J. Plast. 15 (1999) 119–138.

[16] E. Lorentz, A. Benallal, Gradient constitutive relations: numerical aspects and application to gradient damage, Comput. Methods Appl. Mech. Engrg. 194 (2005) 5191–5220.

[17] R.H.J. Peerlings, Enhanced damage modelling for fracture and fatigue, Ph.D. Thesis, Eindhoven University of Technology, The Netherlands, 1999, pp. 65–66.

[18] T. Liebe, P. Steinman, A. Benallal, Theoretical and computational aspects of a thermodynamically consistent framework for geometrically linear gradient damage, Comput. Methods Appl. Mech. Engrg. 190 (2001) 6555–6576.

[19] T. Svedberg, K. Runesson, An algorithm for gradient-regularized plasticity coupled to damage based on a dual mixed FE-formulation, Comput. Methods Appl. Mech. Engrg. 161 (1998) 49–65.

[20] B.J. Dimitrijevic, K. Hackl, A method for gradient enhancement of continuum damage models, Tech. Mechanik 28 (2008) 43–52.

[21] B. Bourdin, G. Francfort, J.-J. Marigo, The Variational Approach to Fracture, Springer, 2008.

[22] B. Halphen, Q.S. Nguyen, Sur les matériaux standard generalises, J. de Mécanique 14 (1975) 39–63.

[23] P. Germain, Q.S. Nguyen, P. Suquet, Continuum thermodynamics, J. Appl. Mech. 50 (1983) 1010–1020.

[24] I. Ekeland, R. Temam, Analyse Convexe et problèmes Variationnels, Dunod, Gauthier-Villars, Paris, 1974.

[25] A. Benallal, J.-J. Marigo, Bifurcation and stability issues in gradient theories with softening, Model. Simul. Mater. Sci. Engrg. 15 (2007) 283–295.

[26] G. Francfort, A. Mielke, Existence results for a class of rate-independent material models with nonconvex elastic energies, J. Reine Angew. Math. 595 (2006) 55–91.

[27] M. Fortin, R. Glowinski, Augmented Lagrangian methods: application to the numerical solution of boundary-value problems, Studies in Mathematics and its Applications, vol. 15, North-Holland, 1983.

[28] A. Ern, S. Meunier, A posteriori error analysis of Euler–Galerkin approximations to coupled elliptic-parabolic problems, ESAIM Math. Mod. Numer. Anal. 43 (2009) 353–375.

[29] A. Simone, H. Askes, R.H.J. Peerlings, L.J. Sluys, Interpolation requirements for implicit gradient-enhanced continuum damage models, Commun. Numer. Methods Engrg. 19 (2003) 563–572.

[30] J.-J. Marigo, Formulation d'une loi d'endommagement d'un matériau élastique, C.R. Acad. Sci. Paris, série IIb 292 (1981) 1309–1312.

[31] P.-B. Badel, V. Godard, J.-B. Leblond, Application of some anisotropic damage model to the prediction of the failure of some complex industrial concrete structure, Int. J. Solids Struct. 44 (2007) 5848–5874.

[32] R. de Borst, Computation of post-bifurcation and post-failure behaviour of strain-softening solids, Comput. Struct. 25 (1987) 211–224.

[33] M.A. Crisfield, Snap-through and snap-back response in concrete structures and the dangers of under-integration, Int. J. Num. Methods Engrg. 22 (1986) 751–767.

[34] E. Lorentz, P. Badel, A new path-following constraint for strain-softening finite element simulations, Int. J. Num. Methods Engrg. 60 (2004) 499–526.

[35] Code_Aster, finite element software distributed by EDF, open-source and freely available on <www.code-aster.org>.

[36] J.R. Rice, The mechanics of earthquake rupture, in: Physics of the Earth interior, in: A.M Dziewonski, E. Boschi (Eds.), Proceedings of International school of physics "Enrico Fermi" course 78, Italian Physical Society and North-Holland Publ., 1979, pp. 555–649.

[37] G. Meschke, P. Dumstorff, Energy-based modeling of cohesive and cohesionless cracks via X-FEM, Comput. Methods Appl. Mech. Engrg. 196 (2007) 2338–2357.

[38] A.S. Kobayashi, N.M. Hawkins, D.B. Barker, B.M. Liaw, Fracture process zone of concrete, in: S.P. Shah (Ed.), Application of fracture mechanics to cementious composites, NATO-ARW, Northwestern University, 1984, pp. 25–50.