Modeling and simulation of quasi-brittle failure with continuous anisotropic stress-based gradient-enhanced damage models

Vandoren, B.; Simone, A.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Accepted Manuscript

Modeling and simulation of quasi-brittle failure with continuous anisotropic stress-based gradient-enhanced damage models

B. Vandoren, A. Simone

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Please cite this article as:
Modeling and simulation of quasi-brittle failure with continuous anisotropic stress-based gradient-enhanced damage models
Vandoren, B. and Simone, A.
Computer Methods in Applied Mechanics and Engineering, 332:644−685, 2018

Modeling and simulation of quasi-brittle failure with continuous anisotropic stress-based gradient-enhanced damage models

B. Vandoren[a,*], A. Simone[b,c]

[a]*Faculty of Engineering Technology, Hasselt University, Hasselt, Belgium*
[b]*Department of Industrial Engineering, University of Padova, Padua, Italy*
[c]*Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands*

Abstract

Two anisotropic stress-based gradient-enhanced damage models are proposed to address the issue of spurious damage growth typical of continuous standard gradient-enhanced damage models. Both models are based on a decreasing interaction length upon decreasing stresses and do not require additional model parameters or extra degrees of freedom when compared to standard gradient-enhanced models. It is observed that with the proposed models damage spreading is significantly reduced due to the occurrence of non-physical oscillations in the nonlocal strain field near the strain localization band. Model improvements to eliminate these strain oscillations upon vanishing length scale values are proposed. The capability of the models and their patched versions to correctly simulate damage initiation and propagation is investigated by means of mode-I failure, shear band and four-point bending tests.

Keywords: gradient-enhanced damage, anisotropic damage, quasi-brittle failure, transient length scale

1. Introduction

Using a nonlocal model for modeling damage in quasi-brittle materials, either in its original integral form [1] or its differential, gradient-enhanced form [2], solves the well-known mesh-dependency issue of local damage models where strains tend to localize in the smallest element of the discretization [3]. However, it has been reported that these nonlocal models may introduce other issues such as an initiation of damage at a wrong location and the artificial spreading of damage [4, 5]. In this paper, we present a class of gradient-enhanced damage models that addresses these unwanted phenomena by employing an anisotropic, stress-dependent nonlocal interaction kernel.

The key feature of the proposed models lies in the use of a transient length scale and thus a decreasing nonlocal activity when stresses decrease. The idea of using a transient length scale or, equivalently, a transient gradient activity parameter in gradient-enhanced damage models was already proposed by Geers et al. [4] who employed a strain-based formulation requiring an additional set of degrees of freedom compared to the standard model (their model was reformulated

---

by Saroukhani et al. [6] to avoid the extra set of degrees of freedom). Geers et al. [7] also investigated the use of a damage-based length scale, a proposition that was picked up very recently by Poh and Sun [8], who developed a phenomenological model motivated by the micromorphic theory. The idea of incorporating a transient length scale in integral nonlocal models was proposed by Pijaudier-Cabot et al. [9] and, more recently, by Nguyen [10] with a mixed local-nonlocal formulation. All these transient nonlocal damage models use an isotropic nonlocal interaction domain.

Bažant [11] proposed an integral nonlocal model with an anisotropic interaction kernel with constant length scale in which the principal stress directions govern the shape of the nonlocal averaging domain. Bažant's model, physically motivated by the fact that nonlocal microcrack interactions occur according to the direction of the dominant microcrack, was subsequently refined by Giry et al. [12] by including a transient length scale, allowing for nonlocal interactions to cancel when stresses decrease. A gradient-enhanced variant of this principal stress-based model, presented in Section 3, shares the same properties as the integral version, as demonstrated by the numerical examples in Section 5: damage initiates at the correct location and its artificial spreading is significantly reduced; unphysical oscillations may however arise when the length scale in one ore more directions becomes too small with respect to the size of the discretization. As investigated in their article, Giry et al. [12] indicated that this phenomenon can be mitigated by means of a lower bound for the length scale to keep nonlocal interaction at a minimum (this is actually done by adopting a minimum value of the length scale related to the discretization).

For the transient gradient-enhanced damage models, the use and definition of such a minimum length scale (or its maximum value to avoid nonlocal interactions) or, more specifically, its relation with the finite element discretization, has received little to no attention. Only Saroukhani et al. [6] mentioned that to avoid nonlocal interactions in the isotropic strain-based model at the beginning of a simulation it was empirically found that a gradient parameter smaller than the square of the distance between any two integration points should be used. In Section 4 of this work, we address the use of such a residual length scale in more detail and derive element size-specific lower bounds that correspond to an *exact* 'local mode' of the gradient-enhanced damage model. It will be shown that adopting values below these lower bounds will, analogously to the nonlocal integral model, result in mesh-dependent solution field oscillations that, although of unphysical nature, may help in delaying spurious damage growth by 'diking' the main localization band. However, even when adopting this numerical lower bound on the length scale, the principal stress-based model still suffers from tortuous force-displacement responses for shear-dominant failure problems, an issue also observed in the nonlocal integral version of the model [12]. An improved anisotropic stress-based gradient-enhanced model is therefore proposed here to solve this issue: an equivalent stress-based model in which the degree of nonlocal interaction does not depend on the magnitude of the principal stresses but depends on an equivalent stress measure linked to the employed material model.

In both anisotropic stress-based gradient models, anisotropy is incorporated by means of a tensorial gradient activity parameter instead of a scalar parameter, an idea also exploited by Wu et al. [13] who employed an anisotropic but constant definition of the interaction kernel to account for the underlying microstructure of composites. For the sake of readability, the mathematical formulation and algorithmic treatment (finite element discretization and consistent linearization) of both models are reported in the appendix.

2

2. Local and isotropic gradient-enhanced damage models

2.1. Local damage models

The concept of damage mechanics provides a phenomenological description of stiffness degradation in softening materials through the constitutive relation

$$\boldsymbol{\sigma} = \mathbf{C}^{\mathrm{s}} : \boldsymbol{\varepsilon} \tag{1}$$

between stresses $\boldsymbol{\sigma}$ and strains $\boldsymbol{\varepsilon}$, where $\mathbf{C}^{\mathrm{s}}$ is a secant stiffness tensor that depends on one or more scalar damage variables ranging between zero (no damage) and one (full damage). The simplest damage-based constitutive relation,

$$\boldsymbol{\sigma} = (1 - \omega)\,\mathbf{C}^{\mathrm{e}} : \boldsymbol{\varepsilon} \tag{2}$$

where $\mathbf{C}^{\mathrm{e}}$ is the elastic stiffness tensor, involves only one damage variable ω and represents the isotropic case. More complex tensorial damage models, such as anisotropic models characterized by a directional dependence of damage evolution in which, e.g., the shear stiffness degrades less rapidly than the normal stiffness, may involve multiple damage variables ω_i.

In essence, most damage models are characterized by two main ingredients: a scalar measure for the strain state (i.e., an equivalent strain $\varepsilon_{\mathrm{eq}}$) and a damage evolution law that governs the decay rate of the load-carrying capacity of the material. In this work, the equivalent strain is described by the modified von Mises model

$$\varepsilon_{\mathrm{eq}} = \frac{k - 1}{2k\,(1 - 2\nu)}I_{1,\varepsilon} + \frac{1}{2k}\sqrt{\left(\frac{k - 1}{1 - 2\nu}I_{1,\varepsilon}\right)^2 + \frac{12k}{(1 + \nu)^2}J_{2,\varepsilon}} \tag{3}$$

in which the first strain invariant $I_{1,\varepsilon}$ is calculated according to

$$I_{1,\varepsilon} = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\,, \tag{4}$$

and the second deviatoric strain invariant $J_{2,\varepsilon}$ is expressed as

$$J_{2,\varepsilon} = \frac{1}{6}\left(\left(\varepsilon_{xx} - \varepsilon_{yy}\right)^2 + \left(\varepsilon_{yy} - \varepsilon_{zz}\right)^2 + \left(\varepsilon_{zz} - \varepsilon_{xx}\right)^2\right) + \varepsilon_{xy}^2 + \varepsilon_{yz}^2 + \varepsilon_{zx}^2 \tag{5}$$

where $\varepsilon_{xy} = \gamma_{xy}/2$, $\varepsilon_{yz} = \gamma_{yz}/2$, and $\varepsilon_{zx} = \gamma_{zx}/2$. It is noted that for plane stress conditions $\varepsilon_{zz} = -\nu\left(\varepsilon_{xx} + \varepsilon_{yy}\right)/(1 - \nu)$. In (3), ν is the Poisson's ratio and the parameter $k = f_{\mathrm{c}}/f_{\mathrm{t}}$ is defined as the ratio between the compressive strength f_{c} and tensile strength f_{t} of the material. Damage growth is described by an evolution law that, in this contribution, takes the form of the exponential law

$$\omega = 1 - \frac{\kappa_0}{\kappa}\left(1 - \alpha + \alpha e^{-\beta(\kappa - \kappa_0)}\right) \tag{6}$$

in which the history parameter κ stores the largest value of the local equivalent strain $\varepsilon_{\mathrm{eq}}$ the material (i.e., an integration point) has ever experienced. The material parameters α and β in (6) are calibrated by experimental tests and define, respectively, the residual stress of the softening branch of the stress-strain curve and the fracture toughness. Damage initiates as soon as the threshold strain value κ_0 has been exceeded and grows according to the Kuhn–Tucker conditions

$$f \le 0\,, \qquad \dot{\kappa} \ge 0\,, \qquad f\dot{\kappa} = 0\,, \tag{7}$$

where the damage loading function

$$f = \varepsilon_{\text{eq}} - \kappa \tag{8}$$

verifies if the material undergoes loading ($f > 0$) or unloading ($f < 0$). In the numerical examples in this contribution, the damage threshold

$$\kappa_0 = \frac{f_{\text{t}}}{E}, \tag{9}$$

where E is the Young's modulus of the material.

2.2. Isotropic gradient-enhanced damage models with a constant length scale

From the formulation of the constitutive model described in the previous section, it is obvious that this model acts *locally*, i.e., at integration point level and without any interaction among neighboring integration points. Nonlocal averaging techniques can be used to avoid the well-known dependence of the numerical solution from the spatial discretization [3]. The gradient-enhanced damage model proposed by Peerlings et al. [2] can be considered as an approximated differential formulation of the nonlocal damage model by Pijaudier-Cabot and Bažant [1]. This differential formulation regularizes the solution field by defining a nonlocal equivalent strain in the implicit expression

$$\tilde{\varepsilon}_{\text{eq}} - c\nabla^2\tilde{\varepsilon}_{\text{eq}} = \varepsilon_{\text{eq}} \tag{10}$$

in which the gradient parameter $c = \frac{1}{2}l^2$, function of the internal length scale l, governs the size of the isotropic, spherical averaging (or interaction) zone, and ∇^2 is the Laplacian operator. The nonlocal equivalent strain $\tilde{\varepsilon}_{\text{eq}}$ now acts as the damage-driving parameter and thus replaces the local equivalent strain ε_{eq} in the damage loading function (8):

$$f = \tilde{\varepsilon}_{\text{eq}} - \kappa. \tag{11}$$

Nonlocal interaction—in fact: diffusion the former term will also be used in this paper when referring to diffusion—takes place at the global (structural) level, in which the diffusion equation (10) and equilibrium equation (neglecting body forces)

$$\nabla \cdot \boldsymbol{\sigma} = \mathbf{0}, \tag{12}$$

with $\nabla\cdot$ denoting the divergence vector operator, are simultaneously solved [2]. This model will be referred to as the 'standard model' in the remainder of this paper.

2.3. Isotropic gradient-enhanced damage models with a transient length scale

It was observed that gradient-enhanced damage formulations with a constant gradient parameter c exhibit spurious damage growth in the final stages of failure [4, 5]. To alleviate this issue, Geers et al. [4] proposed the use of the transient gradient parameter

$$\zeta\left(\varepsilon_{\text{eq}}\right) = \begin{cases} c\left(\frac{\varepsilon_{\text{eq}}}{\varepsilon_{\text{eq},\zeta}}\right)^{n_\zeta} & \text{if } \varepsilon_{\text{eq}} \leq \varepsilon_{\text{eq},\zeta} \\ c & \text{if } \varepsilon_{\text{eq}} > \varepsilon_{\text{eq},\zeta} \end{cases}, \tag{13}$$

that replaces c in the diffusion equation (10). In (13), $\varepsilon_{\text{eq},\zeta}$ and n_ζ are model parameters. Since this strain-based transient-gradient damage model requires an additional continuity equation, and

4

consequently an extra degree of freedom per node, Saroukhani et al. [6] proposed a simplified implementation of the model by dividing the transient diffusion equation by $\zeta \neq 0$, resulting in

$$\frac{\tilde{\varepsilon}_{\text{eq}}}{\zeta} - \nabla^2 \tilde{\varepsilon}_{\text{eq}} = \frac{\varepsilon_{\text{eq}}}{\zeta} \, . \tag{14}$$

As an alternative to strain-based models, the degree of nonlocality can also be related to the stress level, as will be investigated in the next section, or the amount of damage ω, either directly [7] or using a damage-based parameter inspired by micromorphics [8]. In the latter model, the second term in (10) is multiplied by the scaling parameter

$$s = \frac{(1 - R) \exp(-\eta \omega) + R - \exp(-\eta)}{1 - \exp(-\eta)} \, , \tag{15}$$

where η is a model parameter and R allows for residual nonlocal interaction when the damage parameter approaches unity. In these transient models, the averaging (i.e., gradient) term becomes less important upon damage accumulation, similar to the more recent phase-field models for cohesive fracture [14], leading to physically more acceptable damage profiles when compared to models with a constant length scale.

3. Anisotropic stress-based gradient-enhanced models

As an alternative to a strain-based transient gradient-enhanced damage approach, a stress-based model can be formulated in line with the stress-based nonlocal damage model proposed by Giry et al. [12]. In contrast with the models mentioned in the previous section, the spatial diffusive behavior is no longer isotropic but governed by an anisotropic interaction kernel that depends on the stress state at a given material point. Anisotropy is introduced in the gradient-enhanced damage model by means of the anisotropic generalization of the original diffusion equation (10) (modified Helmholtz's equation):

$$\tilde{\varepsilon}_{\text{eq}} - \nabla \cdot \left(\mathbf{c} \nabla \tilde{\varepsilon}_{\text{eq}} \right) = \varepsilon_{\text{eq}} \, . \tag{16}$$

The original scalar gradient parameter c is thus replaced by a second order tensor $\mathbf{c}$ that is, in a two-dimensional framework, represented by the symmetric anisotropic gradient matrix

$$\mathbf{c} = \begin{bmatrix} c_{xx} & c_{xy} \\ c_{xy} & c_{xy} \end{bmatrix} \, . \tag{17}$$

Nonlocal averaging is no longer performed over an isotropic, circular area but according to an ellipse with size, shape, and orientation governed by the components of $\mathbf{c}$. This can be demonstrated by examining the nonlocal equivalent strain fields $\tilde{\varepsilon}_{\text{eq}}$ of the radially-loaded perforated disc in Figure 1. Under the same loading conditions, a change of the gradient parameters leads to an anisotropic localization behavior of the nonlocal equivalent strains. When using the modified diffusion equation (16), the isotropic case can be retrieved by setting

$$\mathbf{c} = \mathbf{c}_c = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \, . \tag{18}$$

In this contribution, two anisotropic models, based on principal stresses and nonlocal equivalent stress, are proposed. Other definitions of the anisotropic length scale tensor $\mathbf{c}$ can also be devised, e.g., based on the microstructure of the material [15].
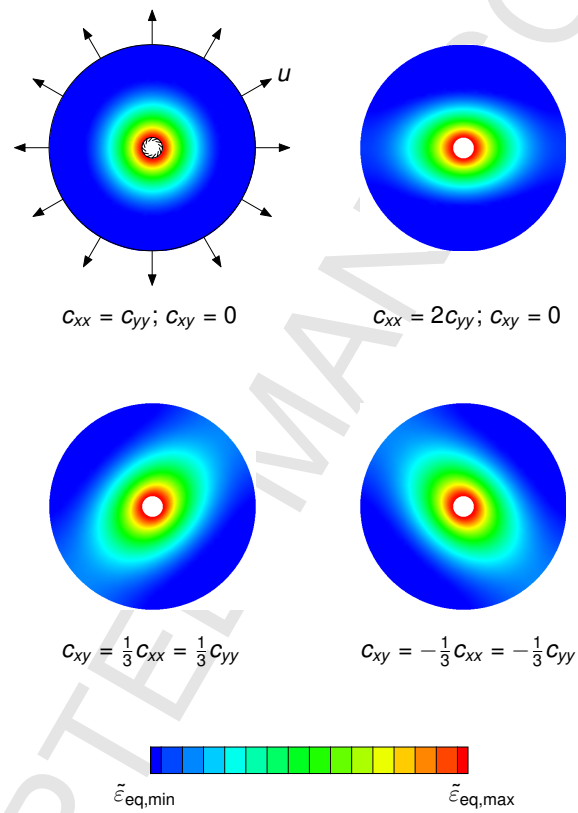
5

Figure 1: Qualitative representation of the influence of the components of the gradient matrix **c** on the nonlocal equivalent strains in a radially-loaded disc. The inner edge of the disc has been constrained whereas a radial displacement u has been applied to the outer edge.

3.1. Anisotropic model based on principal stresses

In this model, which can be regarded as the differential nonlocal counterpart of the integral nonlocal anisotropic model proposed by Giry et al. [12], the anisotropic gradient matrix $\mathbf{c}$ is expressed as

$$\mathbf{c} = \mathbf{c}_\sigma = \mathbf{R} \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} \mathbf{R}^\mathrm{T} \tag{19}$$

in which

$$c_1 = c \left(\frac{\sigma_1}{f_\mathrm{t}} \right)^2 \qquad \text{and} \qquad c_2 = c \left(\frac{\sigma_2}{f_\mathrm{t}} \right)^2 \tag{20}$$

are function of the principal stresses

$$\sigma_{1,2} = \frac{\sigma_{xx} + \sigma_{yy}}{2} \pm \sqrt{\left(\frac{\sigma_{xx} - \sigma_{yy}}{2} \right)^2 + \tau_{xy}{}^2} \,. \tag{21}$$

The rotation matrix

$$\mathbf{R} = \begin{bmatrix} \cos\theta_\mathrm{p} & -\sin\theta_\mathrm{p} \\ \sin\theta_\mathrm{p} & \cos\theta_\mathrm{p} \end{bmatrix}, \tag{22}$$

and the principal angle

$$\theta_\mathrm{p} = \frac{1}{2} \tan^{-1} \left(\frac{2\tau_{xy}}{\sigma_{xx} - \sigma_{yy}} \right) \,. \tag{23}$$

The components of $\mathbf{c}$ can thus be expressed as

$$c_{xx} = c_1 \cos^2\theta_\mathrm{p} + c_2 \sin^2\theta_\mathrm{p} \,, \tag{24}$$

$$c_{yy} = c_1 \sin^2\theta_\mathrm{p} + c_2 \cos^2\theta_\mathrm{p} \,, \tag{25}$$

and

$$c_{xy} = c_{yx} = (c_1 - c_2) \cos\theta_\mathrm{p} \sin\theta_\mathrm{p} \,. \tag{26}$$

To avoid algorithmic difficulties when calculating principal angles close to 90°, these expressions are rewritten as

$$c_{xx} = \frac{c}{f_\mathrm{t}^2} \left(\sigma_{xx}^2 + \tau_{xy}^2 \right) \,, \tag{27}$$

$$c_{yy} = \frac{c}{f_\mathrm{t}^2} \left(\sigma_{yy}^2 + \tau_{xy}^2 \right) \,, \tag{28}$$

and

$$c_{xy} = \frac{c}{f_\mathrm{t}^2} \tau_{xy} \left(\sigma_{xx} + \sigma_{yy} \right) \tag{29}$$

as demonstrated in Appendix A.

7

3.2. Anisotropic model based on equivalent stress

A characteristic of the anisotropic model presented in the previous section is that the size of the averaging volume is independent of whether there are tensile or compressive stresses acting at a certain material point. For example, $c_{xx} = c$ and $c_{yy} = c_{xy} = 0$ mm^2 for these two uniaxial stress states: $\{\sigma_{xx}, \sigma_{yy}, \tau_{xy}\} = \{f_t, 0, 0\}$ N/mm^2 and $\{\sigma_{xx}, \sigma_{yy}, \tau_{xy}\} = \{-f_t, 0, 0\}$ N/mm^2. This implies that, even if the material model (3) does penalize tensile states more than compressive states via the factor k, the diffusion behavior makes no distinction between compressive and tensile stresses. Therefore, a new definition of the anisotropic gradient matrix

$$\mathbf{c} = \mathbf{c}_{\tilde{\sigma}} = s\,c\,\mathbf{c}_{\sigma,\text{norm}} \tag{30}$$

is proposed in which $s = \tilde{\sigma}_{\text{eq}}/f_t$ is a scaling factor ranging between zero and one that takes into account the employed material model through the equivalent stress measure $\tilde{\sigma}_{\text{eq}}$, defined next, and $\mathbf{c}_{\sigma,\text{norm}}$ is the normalized $\mathbf{c}_\sigma$ from the principal stress-based model (Equation (19)). The largest singular value norm has been employed for the evaluation of $\mathbf{c}_{\sigma,\text{norm}}$, i.e., each component is divided by the largest eigenvalue of $\mathbf{c}_\sigma$, which is in this case the product of the square of the largest principal stress and c/f_t^2, resulting in

$$\mathbf{c}_{\sigma,\text{norm}} = \begin{bmatrix} c_{xx,\text{norm}} & c_{xy,\text{norm}} \\ c_{xy,\text{norm}} & c_{yy,\text{norm}} \end{bmatrix} \tag{31}$$

with

$$c_{xx,\text{norm}} = \frac{\sigma_{xx}^2 + \tau_{xy}^2}{\max\left(\sigma_1{}^2, \sigma_2{}^2\right)}\,, \tag{32}$$

$$c_{yy,\text{norm}} = \frac{\sigma_{yy}^2 + \tau_{xy}^2}{\max\left(\sigma_1{}^2, \sigma_2{}^2\right)}\,, \tag{33}$$

and

$$c_{xy,\text{norm}} = \frac{\tau_{xy}\left(\sigma_{xx} + \sigma_{yy}\right)}{\max\left(\sigma_1{}^2, \sigma_2{}^2\right)}\,. \tag{34}$$

When comparing the definitions of $\mathbf{c}_\sigma$ and $\mathbf{c}_{\tilde{\sigma}}$, it can be noticed that both tensors (and the stress tensor $\boldsymbol{\sigma}$) are coaxial, and $\mathbf{c}_\sigma$ and $\mathbf{c}_{\tilde{\sigma}}$ are a scalar multiple of each other through the relation

$$\mathbf{c}_{\tilde{\sigma}} = \frac{\tilde{\sigma}_{\text{eq}} f_t}{\max\left(\sigma_1{}^2, \sigma_2{}^2\right)} \mathbf{c}_\sigma\,. \tag{35}$$

The meaning of the definition of $\mathbf{c}_{\tilde{\sigma}}$ (30) can be described as follows: in this model the anisotropy, i.e., the *shape* of the interaction kernel, is governed by the stress state through $\mathbf{c}_{\sigma,\text{norm}}$, whereas the *size* of the interaction kernel is governed by the ratio between nonlocal equivalent stress and tensile strength through s. The matrix $\mathbf{c}_{\sigma,\text{norm}}$ can thus be considered as a 'unit ellipse' that is scaled by $s = \tilde{\sigma}_{\text{eq}}/f_t$ in which the nonlocal equivalent stress

$$\tilde{\sigma}_{\text{eq}} = (1 - \omega)\,E\tilde{\varepsilon}_{\text{eq}} \tag{36}$$

is energetically conjugate to the nonlocal equivalent strain $\tilde{\varepsilon}_{\text{eq}}$. Due to the nature of the constitutive behavior (7), $\tilde{\sigma}_{\text{eq}}$ in (36) will never exceed the tensile strength f_t. Indeed, by recalling the definition of the damage threshold (9), the scalar s can also be rewritten as

$$s = (1 - \omega)\,\frac{\tilde{\varepsilon}_{\text{eq}}}{\kappa_0}\,. \tag{37}$$
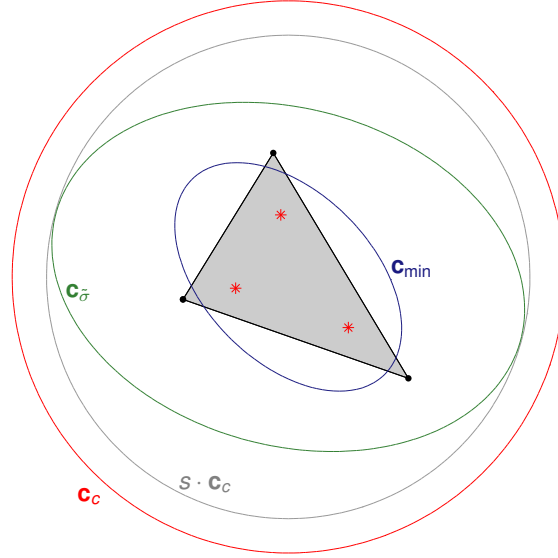
8

Figure 2: Graphical representation of the anisotropic nonlocal equivalent stress-based damage model. The outer circle (in red, denoted by $\mathbf{c}_c$) represents the isotropic interaction kernel of the standard model, expressed by (18). This averaging volume is subsequently scaled by the factor s (circle in gray). Finally, anisotropy is introduced by incorporating the stress state (Equation (30)), resulting in the interaction volume denoted by $\mathbf{c}_{\tilde{\sigma}}$ (in green). The ellipse in blue ($\mathbf{c}_{\min}$) represents the lower bound of the averaging volume that guarantees non-oscillatory nonlocal behavior for a given triangular element (with integration points marked by red asterisks) as discussed in Section 4.

in which the product $(1 - \omega)\,\tilde{\varepsilon}_{\mathrm{eq}}$ will never be larger than the damage threshold κ_0.

Figure 2 provides a graphical representation of this model, where $\mathbf{c}_c$ denotes the isotropic interaction volume governed by the gradient parameter c in Equation (18) that is subsequently scaled by the scaling factor s in Equation(37) (gray circle) upon a decreasing nonlocal equivalent stress. The inscribed green ellipse, tagged by $\mathbf{c}_{\tilde{\sigma}}$, represents the final, anisotropic averaging volume, calculated according to (30). The ellipse tagged by $\mathbf{c}_{\min}$ indicates the lower bound of the interaction volume that guarantees nonlocal behavior for a given triangular finite element, as will be further discussed in Section 4.

The proposed definition of the gradient matrix (30) offers a more general framework to incorporate anisotropy in existing transient isotropic gradient-enhanced damage models. Indeed, the scaling factor s can also be related to other quantities, such as a damage parameter derived from micromorphics [8] (Equation (15)) or other phenomenological approaches [7], as discussed in Section 2.3. Note however that a damage evolution law with residual stress, such as the exponential law (6) with $\alpha < 1$, necessitates a dedicated evolution law for the scaling factor s to completely cancel nonlocal interactions upon strain localization. As it can be observed when combining Equations (6) and (37), s will approach $1 - \alpha$ when strains localize, still allowing nonlocal interactions since the gradient matrix $\mathbf{c}$ does not vanish from the diffusion equation (16). To overcome this issue in a simple way, the scalar damage parameter ω in (37) can be replaced
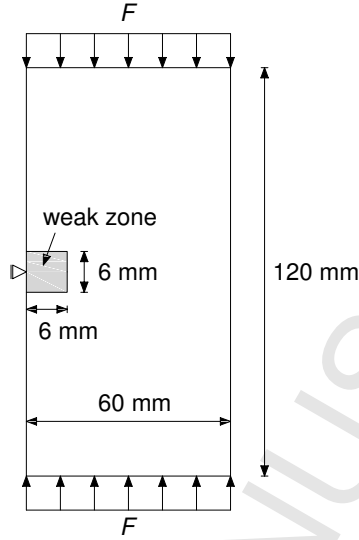
9

Figure 3: Shear band test: Geometry and boundary conditions. The weak zone contains elements with a reduced damage threshold.

by an alternative scalar ω_s that obeys its separate evolution law

$$\omega_s = 1 - \frac{\kappa_0}{\kappa} e^{-\beta_s(\kappa - \kappa_0)} , \tag{38}$$

obtained by setting $\alpha = 1$ in the damage evolution law (6). The parameter β_s governs the decay rate of the amount of nonlocal interactions.

Both stress-based models are discretized and linearized using standard techniques as shown in Appendix B.

3.3. Model performance: Shear band problem

The performance of the stress-based models is evaluated through the simulation of a rectangular plate subjected to a compressive load (Figure 3). The plate is discretized using three structured meshes containing 7200, 11,400 and 16,200 triangular elements (only the top half is modeled). The nonlocal equivalent strain and displacement fields are interpolated using linear (T3) and quadratic (T6) interpolants, respectively, to avoid stress oscillations caused by using the same interpolation order for both fields. Plane strain conditions are assumed, Young's modulus $E = 20,000$ N/mm^2, Poisson's ratio $\nu = 0.2$, and a gradient parameter $c = 2$ mm^2 have been adopted. The softening behavior is governed by the exponential damage evolution law (6) with parameters $\alpha = 0.99$ and $\beta = 100$. Damage initiates after the von Mises equivalent strain, with $k = 1$ in Equation (3), exceeds the threshold $\kappa_0 = 0.0001$. As indicated in Figure 3, a weak zone ($\kappa_0 = 0.00005$) is inserted to introduce damage on the left-hand side. The loading response may be characterized by snap-backs and is therefore traced using an energy-release control algorithm [16].

This problem is often considered to highlight a major flaw of the standard gradient-enhanced damage model: the shear band is not stationary. This is clearly an artifact of the model as experimental evidence (see, e.g., References [17, 18] and references therein) indicates that shear
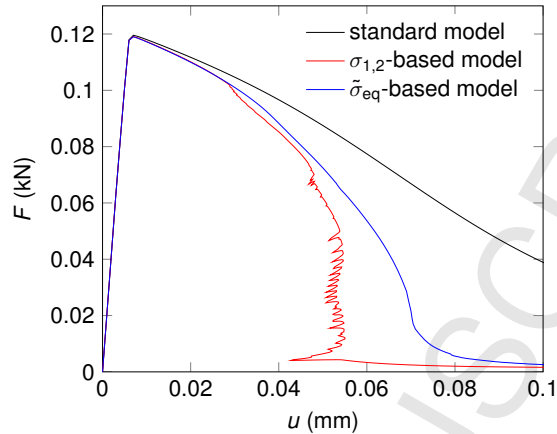
10

Figure 4: Shear band test: Force-displacement curves for the original (standard), principal stress-based ($\sigma_{1,2}$), and nonlocal equivalent stress-based ($\tilde{\sigma}_{\text{eq}}$) model (7200 elements).

bands are stationary: once a shear band forms, the deformation increases in a delimited and stationary band.

Figure 4 shows that the use of the stress-based models results in a more brittle response when compared to the standard model. Since these results are obtained using the same model parameters (α, β, c) and they influence the fracture properties, their recalibration, which is beyond the scope of this work, is deemed necessary when the same structural response is desired. The principal stress-based model gives also rise to a very tortuous force-displacement curve, as evident from Figure 5. This phenomenon was also observed in the stress-based integral nonlocal damage model by Giry et al. [12] (Figure 13 (a) in their manuscript): the larger the mesh size, the smoother the force-displacement curve. As can be noticed from Figure 6, this behavior is absent in case of the nonlocal equivalent stress-based model.

When comparing the evolution of the damage fields in Figure 7, both stress-based models do not suffer from damage spreading, even at high deformations levels (imposed top displacement of 1 mm). The standard model, on the other hand, suffers from unphysical spreading of the damage field, after a shear band has been formed (left column of snapshots in Figure 7). At first sight, the stress-based models thus seem to solve the damage spreading issue encountered with the standard model. However, when inspecting the (normalized) nonlocal equivalent strain fields in Figure 8, which is the quantity that drives damage evolution through the loading function (11), it can be noticed that these fields suffer from unphysical oscillations at higher deformation levels and even include negative values, as can be seen from a slice plot in Figure 9. These negative values will prevent damage from spreading, at least initially, since the damage loading function (11) becomes negative. Damage spreading is thus limited in an artificial way: the oscillations 'dike' the main localization band, reducing the spurious spreading of damage.

The occurrence of this oscillatory behavior, investigated in more detail in the next section, is caused by a vanishing gradient (length scale) parameter. This can also be observed by setting $c = 0$ mm^2 in the standard model (10): the nonlocal equivalent strain fields in Figure 10 suffer from severe oscillations, suggesting that a lower bound has to be defined for the gradient parameter to avoid these unphysical artifacts. The oscillations and corresponding diking effect are thus not only an intrinsic issue of stress-based models: any transient gradient-enhanced model
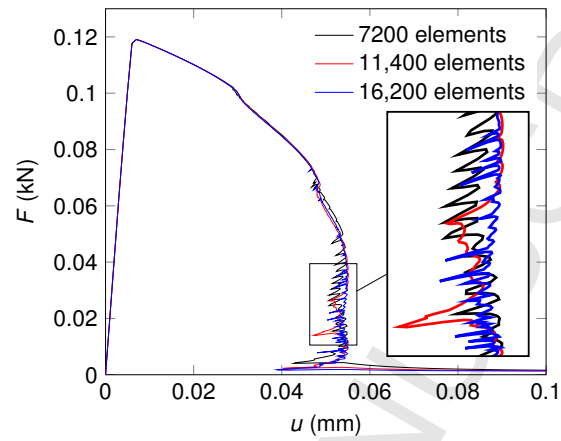
11

Figure 5: Shear band test using the principal stress-based model: Force-displacement curves for various mesh sizes.
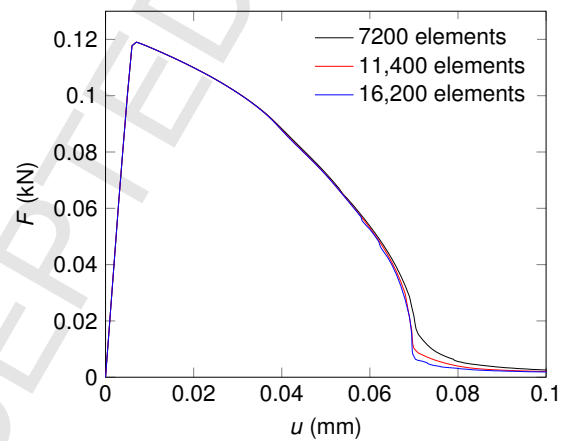

Figure 6: Shear band test using the nonlocal equivalent stress-based model: Force-displacement curves for various mesh sizes.
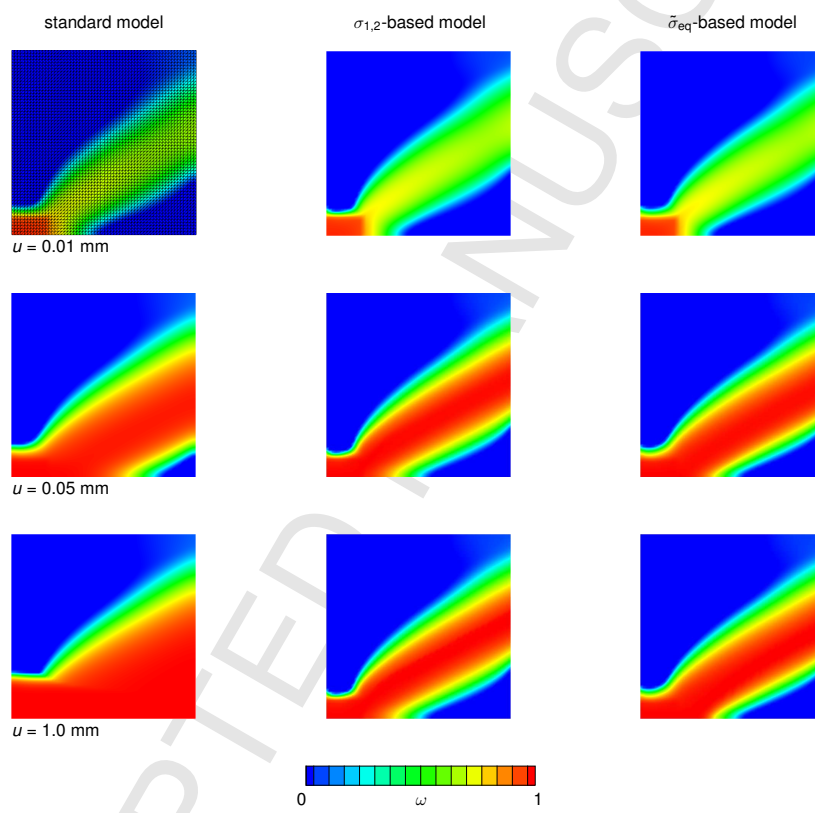
12

Figure 7: Shear band test: Evolution of the damage fields for the original and stress-based models (7200 elements; the first snapshot displays the employed mesh).
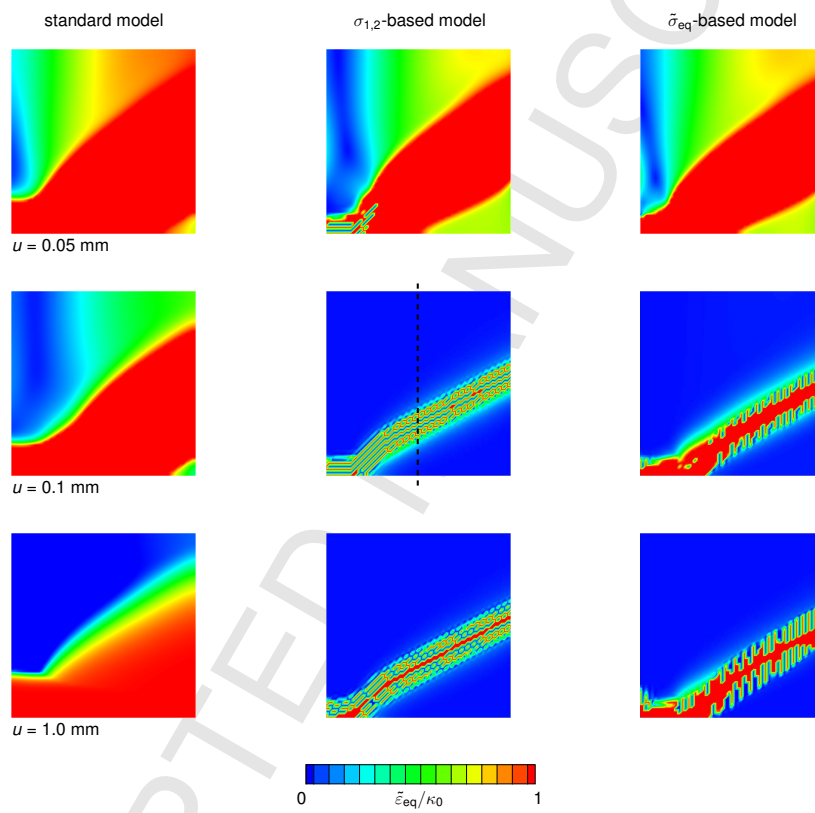
Figure 8: Shear band test: Evolution of the normalized nonlocal equivalent strain fields for the original and stress-based models (7200 elements). Figure 9 shows the normalized nonlocal equivalent strain along the dashed line.
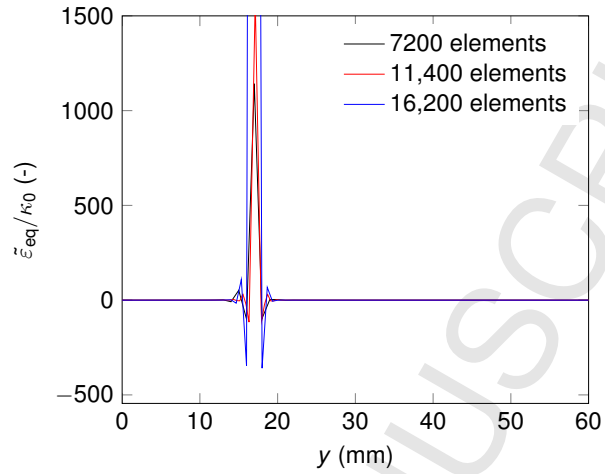
14

Figure 9: Shear band test using the principal stress-based model (top-displacement of $u = 0.1$ mm): Normalized nonlocal equivalent strains along the vertical center line of the specimen.
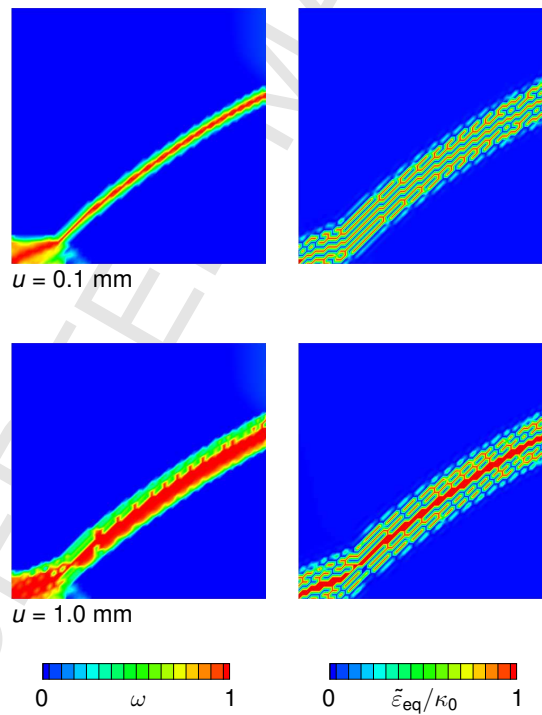

Figure 10: Shear band test using the standard model with $c = 0$ mm^2: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column). The specimen is modeled using 7200 triangular elements.

with a vanishing length scale may also suffer from this behavior, although the severity of the oscillations depends on the chosen model parameters and, evidently, the chosen mesh size. The transient model by Poh and Sun [8], for instance, which can be expressed by (30) with $\mathbf{c}_{\sigma,\text{norm}} = \mathbf{I}$ and a scaling parameter s given by (15), also suffers from this artificial diking effect if the parameters R, which governs the amount of residual nonlocal interaction, c and the mesh resolution are chosen inappropriately. With $R = 0.005$, $c = 2$ mm^2 and a Cartesian mesh consisting of 14,400 quadrilateral elements, the model leads to oscillations but only minor damage spreading, as depicted by Figure 11 and the slice plot in Figure 14. If the gradient parameter is changed to $c = 9$ mm^2, implying a residual gradient parameter $cR = 0.045$ mm^2, the oscillations are absent but spreading of damage is more pronounced as can be noticed from Figure 11. On the other hand, when $c = 9$ mm^2 and $R = 0.001$ (residual gradient parameter equal to 0.009 mm^2), Figures 13 and 14 indicate that oscillations are present again, yet spreading of damage is less pronounced. The force-displacement curves in Figure 15 show the influence of the chosen model parameters on the structural response: when using a gradient parameter $c = 9$ mm^2, a lower residual R results in a more brittle behavior. On the other hand, when the same residual $R = 0.005$ is used, a lower c will, as expected, lead to a more brittle response since damage and consequently the amount of dissipated energy is confined to a smaller volume (this can also be observed when comparing the width of the damage fields in Figures 11 and 12.)

From these observations, it is clear that there exists a relation between the minimum length scale, the employed mesh size and the potential occurrence of oscillations. In the next section, a finite element-size specific lower bound of the gradient parameter is derived that sets the limit between oscillatory and nonlocal smoothing behavior.

4. Oscillation-free stress-based gradient-enhanced damage models

As observed in the previous section, the use of a transient length scale in gradient-enhanced damage models may result in numerical issues when this quantity becomes too small with respect to the mesh size. For nonlocal integral models, this issue was also noted by Giry et al. in their stress-based model [12], as they proposed the use of a minimum length scale equal to the square root of the finite element area to avoid spurious oscillations in the solution fields. To the best of the authors' knowledge, an exact element size-dependent limit on the length scale, which leads to either oscillatory, nonlocal or local behavior, has not been established yet for gradient-enhanced damage models. Only Saroukhani et al. [6] found empirically that, for the strain-based gradient-enhanced model by Geers et al. [4], the gradient parameter should remain larger than the square of the distance between any two integration points to avoid nonlocal interactions at the beginning of a simulation. Next, values for c (for one-dimensional problems) and $\mathbf{c}$ (for two-dimensional problems) are derived that can be considered as a lower bound below which unphysical oscillations take place.

4.1. One-dimensional case

The influence of the choice of the value of c with respect to the mesh size is investigated by observing the localization behavior of a truss loaded by a tensile load (Figure 16). To this end, a rather coarse mesh is used: eight 1.25 mm-long elements in the central part with a reduced cross-section area of 9 mm^2, and 72 elements of equal length elsewhere (cross-section area 10 mm^2). The displacement field is approximated using quadratic shape functions (three nodes per element), whereas the nonlocal equivalent strain field is approximated using linear shape functions
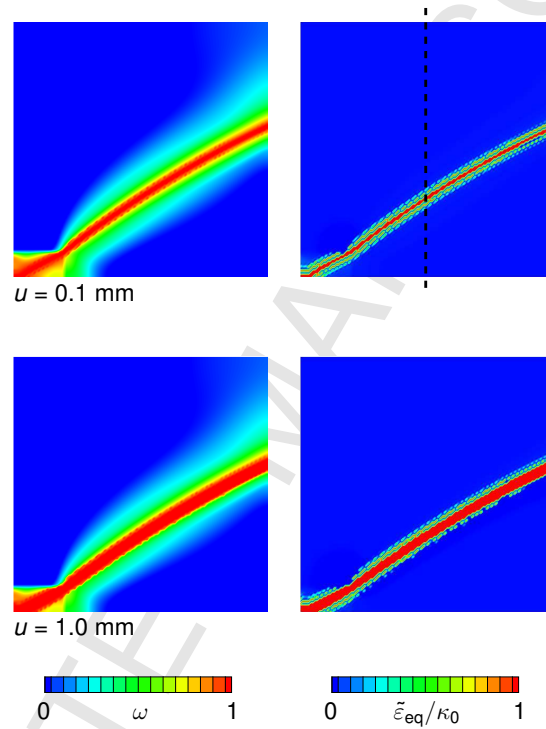
16

Figure 11: Shear band test using the micromorphics-based model by Poh and Sun [8] using model parameters $c = 2$ mm^2 and $R = 0.005$: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column). Figure 14 shows the values along the dashed line. The specimen is modeled using 14,400 quadrilateral elements.
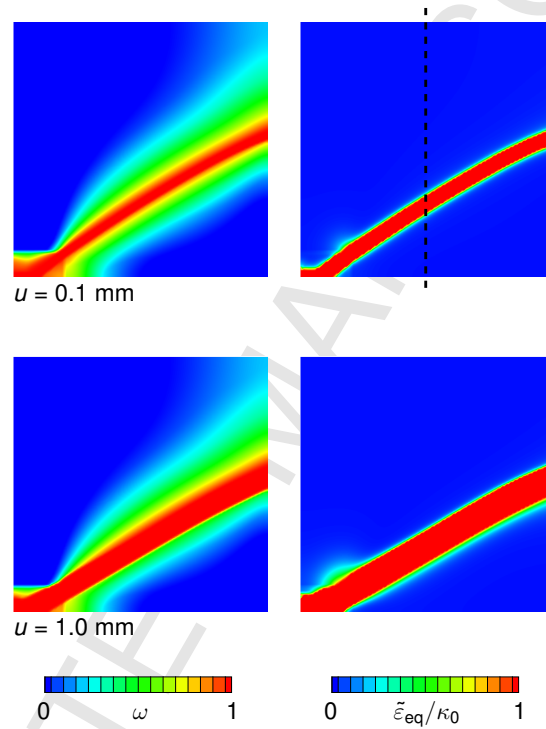
17

Figure 12: Shear band test using the micromorphics-based model by Poh and Sun [8] using model parameters $c = 9$ mm^2 and $R = 0.005$: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column). Figure 14 shows the values along the dashed line. The specimen is modeled using 14,400 quadrilateral elements.
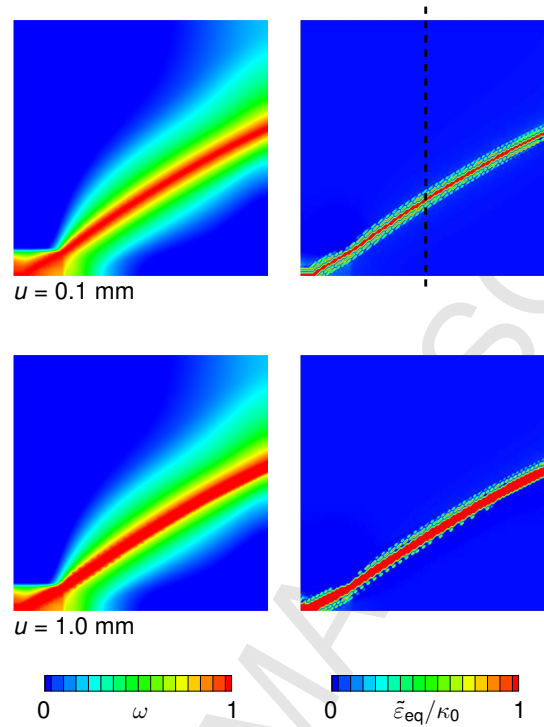
18

Figure 13: Shear band test using the micromorphics-based model by Poh and Sun [8] using model parameters $c = 9$ mm^2 and $R = 0.001$: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column). Figure 14 shows the values along the dashed line. The specimen is modeled using 14,400 quadrilateral elements.
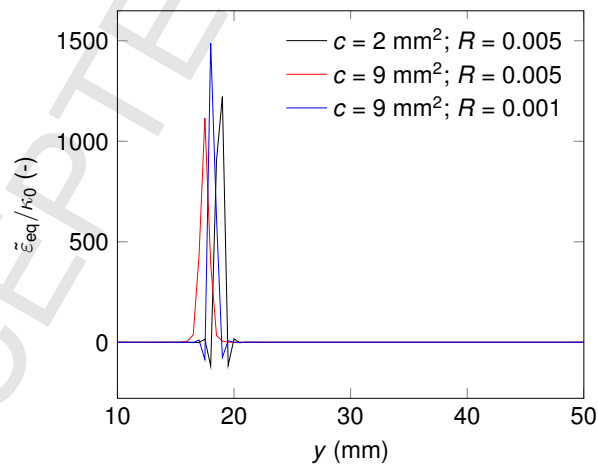

Figure 14: Shear band test using the micromorphics-based model by Poh and Sun [8] (top-displacement of $u = 0.1$ mm): Normalized nonlocal equivalent strains along the vertical center line of the specimen.
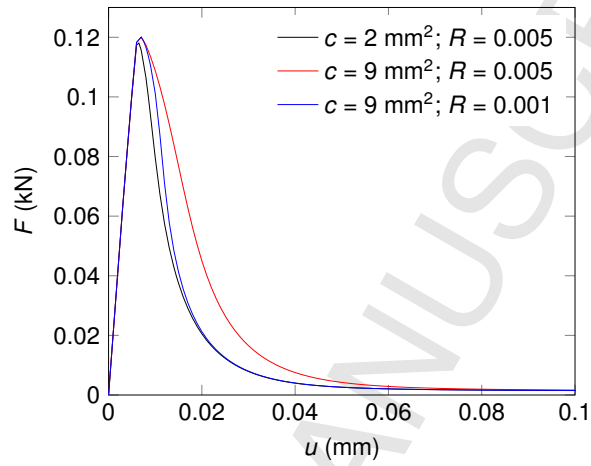
19

Figure 15: Shear band test using the micromorphics-based model by Poh and Sun [8]: Force-displacement curves for three sets of model parameters (14,400 elements).
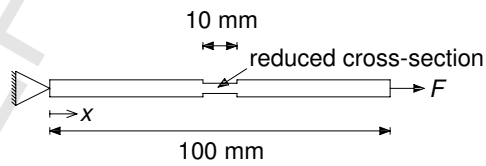

Figure 16: One-dimensional localization test: Geometry and boundary conditions.
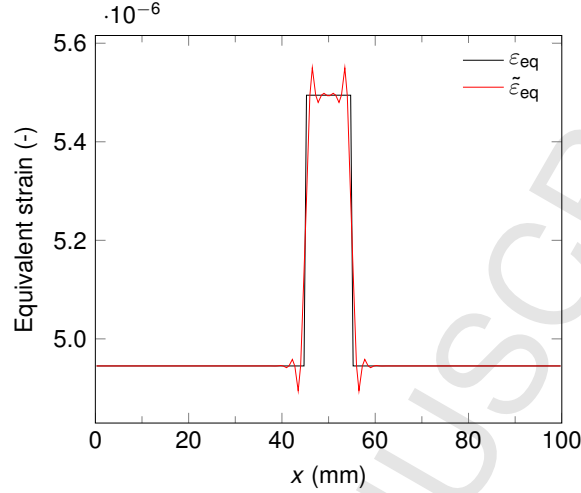
20

Figure 17: One-dimensional localization test (end-displacement of $u = 0.0005$ mm): Overshooting behavior of the nonlocal equivalent strain field when $c = 1$ mm^2.

(two nodes per element). When the nonlocal equivalent strain exceeds the damage threshold $\kappa_0 = 0.0001$, the softening behavior is governed by the linear damage evolution law

$$\omega = 1 - \frac{\kappa_0}{\kappa} \frac{\kappa_c - \kappa}{\kappa_c - \kappa_0} \tag{39}$$

in which $\kappa_c = 0.0125$ denotes the critical value of the nonlocal equivalent strain at which $\omega = 1$. The gradient parameter $c = 1$ mm^2 is scaled according to the one-dimensional version of the principal stress-based model presented in Section 3.1 so that the scaling factor

$$s = \left(\frac{\sigma}{f_t}\right)^2, \tag{40}$$

where the tensile strength $f_t = E\kappa_0$ and σ represents the normal stress in the truss, with the Young's modulus $E = 20{,}000$ N/mm^2.

The bar is loaded using a displacement-controlled incremental-iterative scheme (step size $u = 0.0005$ mm) that switches to energy release control [16] (tolerated dissipated energy per increment $E_{tol} = 0.002$ Nmm) after encountering a snap-back point along the force-displacement curve. The first increment (with end-displacement $u = 0.0005$ mm) corresponds to a normal stress of $\sigma = 0.1099$ N/mm^2 in the part with reduced cross-section and $\sigma = 0.0989$ N/mm^2 elsewhere. Using the definition of the scaling factor s (40), the corresponding scaled gradient parameters c_s are equal to 0.0030 mm^2 and 0.0024 mm^2, respectively. These values appear to be too small for the current discretization since, even though the material is still in its elastic phase, the discontinuous character of the local equivalent strain in combination with a low stress value will lead to an oscillatory overshooting behavior in the nonlocal equivalent strain field as shown in Figure 17. When the same model is used with a gradient parameter $c = 1000$ mm^2, the respective scaled gradient parameters are 3.0 mm^2 and 2.4 mm^2, canceling the oscillations in the nonlocal equivalent strain field as shown in Figure 18. This value of c seems unphysically
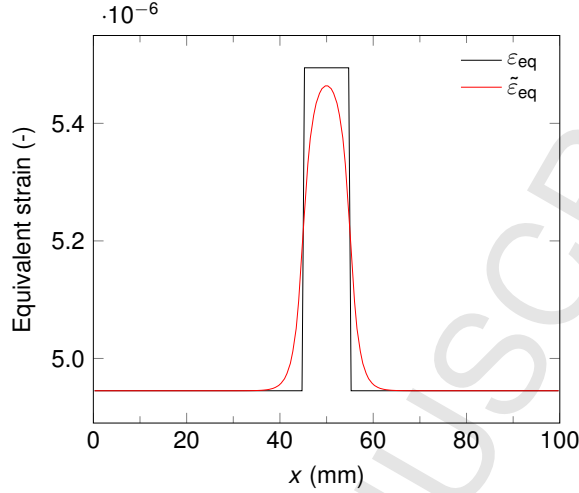
21

Figure 18: One-dimensional localization test (end-displacement of $u = 0.0005$ mm): Correct smoothing behavior of the nonlocal equivalent strain field when $c = 1000$ mm^2.

large since it corresponds to a length scale of 44 mm, almost half the length of the specimen. However, since the scaling factor depends on the square of the normal stress, which is very small in the first load step, the scaled gradient parameter $c_\mathrm{s} = s\,c$ will be small as well.

From a mathematical point of view, the crucial difference between both values of c appears to be in the occurrence of positive off-diagonal terms in the element stiffness matrix

$$\mathbf{K}_{\varepsilon\varepsilon} = \int_\Omega \left(\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon + \mathbf{B}_\varepsilon^\mathrm{T} c_\mathrm{s} \mathbf{B}_\varepsilon \right) \mathrm{d}\Omega \tag{41}$$

that governs the diffusion behavior in which $\mathbf{N}_\varepsilon$ is a matrix containing the finite element shape functions that interpolate the nodal values of the nonlocal equivalent strains, and $\mathbf{B}_\varepsilon = \nabla\mathbf{N}_\varepsilon$ with ∇ the gradient operator (the reader is referred to Appendix B for the derivation of this matrix). As long as these off-diagonal terms are negative, nonlocal smoothing is preserved whereas positive off-diagonal terms lead to oscillations in the solution field of $\tilde{\varepsilon}_\mathrm{eq}$. The turning point between smoothing and oscillatory behavior is thus determined by a value of c_s that leads to zero off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$, yielding therefore a diagonalized or lumped matrix. From a physical point of view, using a diagonalized $\mathbf{K}_{\varepsilon\varepsilon}$ implies that there no longer exists a direct interaction between the nonlocal equivalent strain degrees of freedom (i.e., the element follows a local behavior). This fact illustrates one of the key differences between integral nonlocal and implicit gradient-enhanced damage models discretized using finite elements: in the former type of models, nonlocal interaction occurs at the *integration point* level of the finite elements through nonlocal integral averaging whereas in gradient-enhanced models the nonlocal interaction takes place at the *nodal* (degree of freedom) level through a diffusion equation. It should thus be emphasized that this oscillatory behavior is not a flaw of the model itself but rather a spurious feature caused by the discretization technique.

When inspecting the expression of $\mathbf{K}_{\varepsilon\varepsilon}$ in (41), it can be noticed that the only term that can cause positive off-diagonal entries is $\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon$. More specifically, for a two-node finite element, the shape function matrix $\mathbf{N}_\varepsilon = [0.5(1 - r) \quad 0.5(1 + r)]$ with the local element coordinate $r \in$

22

{−1, 1}, resulting in

$$\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon = 0.25 \begin{bmatrix} (1-r)^2 & 1-r^2 \\ 1-r^2 & (1+r)^2 \end{bmatrix}. \tag{42}$$

This matrix can be considered as a pseudo-mass matrix [19, 20]. Since r^2 will be always smaller than one (unless a nodal integration scheme is employed), the off-diagonal terms remain positive. The other term in $\mathbf{K}_{\varepsilon\varepsilon}$, the diffusivity matrix

$$\mathbf{B}_\varepsilon^\mathrm{T} c_\mathrm{s} \mathbf{B}_\varepsilon = c_\mathrm{s} \begin{bmatrix} l_\mathrm{el}^{-2} & -l_\mathrm{el}^{-2} \\ -l_\mathrm{el}^{-2} & l_\mathrm{el}^{-2} \end{bmatrix}, \tag{43}$$

with l_el the finite element length, will always contain negative off-diagonal terms. To suppress the off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$, thus canceling nonlocal or nodal interaction from a numerical point of view, an exact value of the gradient parameter c_s can be derived that compensates for the positive off-diagonal term in $\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon$:

$$c_\mathrm{min} = \frac{1}{6} l_\mathrm{el}^2, \tag{44}$$

in which a two-point Gauss quadrature is assumed ($r = \pm 1/\sqrt{3}$). Taking into account that $c = \frac{1}{2} l^2$ [2], this minimum value of the gradient parameter corresponds to a minimum (or residual) length scale

$$l_\mathrm{min} = \frac{1}{\sqrt{3}} l_\mathrm{el} \tag{45}$$

that is, coincidentally, equal to the distance between the integration points. It can be concluded that, although the strong form of the implicit gradient-enhanced damage model (10) indicates that local behavior is retrieved when the gradient parameter equals zero, local behavior of the discretized weak form is only recovered with a non-null value of c_s. Using $c_\mathrm{s} = 0$ will indeed cancel the diffusivity matrix $\mathbf{B}_\varepsilon^\mathrm{T} c_\mathrm{s} \mathbf{B}_\varepsilon$ but not the pseudo-mass matrix $\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon$, resulting in positive off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$, causing oscillations in the solution field of $\tilde{\varepsilon}_\mathrm{eq}$. Furthermore, when $c_\mathrm{s} = 0$, the equality $\tilde{\varepsilon}_\mathrm{eq} = \varepsilon_\mathrm{eq}$ is only satisfied in a weak sense (see Equation (B.2)).

It should be noted that a diagonalized system matrix $\mathbf{K}_{\varepsilon\varepsilon}$ can also be obtained by using a reduced two-point nodal integration scheme for the two-node truss elements. This scheme will generate a diagonal pseudo-mass matrix $\mathbf{N}_\varepsilon^\mathrm{T}\mathbf{N}_\varepsilon$ (42) since r equals either 1 or −1. Alternatively, this matrix can be replaced by its diagonally lumped version diag$(0.5(1-r) \quad 0.5(1+r))$. This corresponds to using piecewise constant shape functions that are however no longer consistent with those used to generate $\mathbf{B}_\varepsilon$ in the diffusivity term (43). The use of such an inconsistent pseudo-mass matrix is not preferred here since it may result in a lower rate of convergence with respect to mesh refinement as shown by the blue force-displacement curves in Figure 19: the standard model with a lumped pseudo-mass matrix yields a 'higher' force-displacement curve than the model with consistent pseudo-mass matrix with the same number of elements.

Figure 20 shows the local and nonlocal equivalent strain fields at the first load increment when the lower bound of c_s is limited to c_min in (44). Both fields coincide but at the discontinuity in the cross-section where $\tilde{\varepsilon}_\mathrm{eq}$ is C^0 continuous since it is interpolated using C^0-continuous finite elements. The local strain field is discontinuous because of the C^0 continuity of the displacements, allowing for jumps in the strain field. In other words, when c_s is set to c_min, smoothing or nonlocal averaging does no longer take place at the global level but is restricted to a neighboring-element level: a jump of the local equivalent strain between two adjacent elements will result in an averaged value of the nonlocal equivalent strain field at the element boundary, as indicated
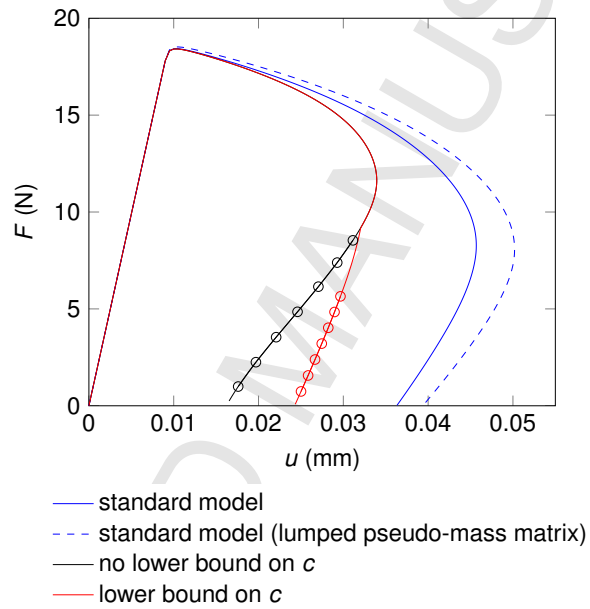
23

Figure 19: One-dimensional localization test using 80 elements: Force-displacement curves for the standard model with consistent (solid blue) and lumped (dashed blue) pseudo-mass matrix, and the stress-based model with (red) and without (black) a lower bound on c_s. The circles refer to the nonlocal equivalent strain fields displayed in Figures 23 and 24.
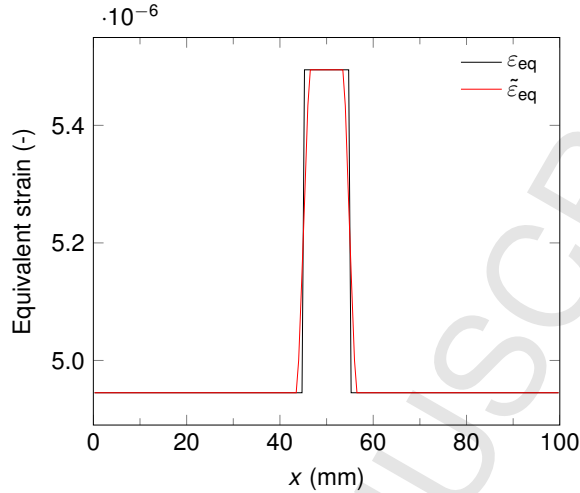
24

Figure 20: One-dimensional localization test (end-displacement $u = 0.0005$ mm): Local behavior of the nonlocal equivalent strain field when the gradient parameter equals c_{min}.

by $\tilde{\varepsilon}_{eq,nodal}$ in the close-up in Figure 21. Since c_{min} is an element-specific value, this phenonenon also holds for non-uniform meshes as demonstrated in Figure 22, in which the number of elements in the part with the reduced cross-section has been doubled. This inherent property of the piecewise continuous nonlocal equivalent strain field may still lead to spreading of damage, even when nonlocal interaction is fully canceled, since a situation may occur where the interpolated value of the nonlocal equivalent strain (indicated by $\tilde{\varepsilon}_{eq,IP}$ in Figure 21), which drives the damage through the loading function (11), becomes higher than the damage threshold κ_0, whereas the local equivalent strain ($\varepsilon_{eq,IP}$ in Figure 21) would not lead to damage since $\varepsilon_{eq,IP} < \kappa_0 < \tilde{\varepsilon}_{eq,IP}$. Indeed, when setting $c_s = c_{min}$ (or $c_s = 0$, as argued above), $\tilde{\varepsilon}_{eq,IP}$ will never match its local counterpart in the elements surrounding a local strain field discontinuity because of the difference in continuity order of both fields, leading to damage initiation in a neighboring element when $\varepsilon_{eq,IP} < \kappa_0 < \tilde{\varepsilon}_{eq,IP}$, This leads to a lower element stiffness in that element, in turn resulting into a higher local equivalent strain and ultimately leading to a higher nonlocal equivalent strain and, again, to the increase of damage. In this way, the element next to a local strain discontinuity 'infects' its neighbor, leading to the spreading of damage, even when nonlocal interaction is canceled. A clear distinction should thus be made between this 'continuity spreading' phenomenon, which is related to the discretization (and interpolation) technique, and 'diffusion spreading', caused by a nonzero value of the gradient parameter c. In a finite element method context, continuity spreading can be postponed by making use of finer meshes or completely canceled by introducing a discontinuity in the nonlocal strain field or in both problem fields as proposed by Simone et al. [21]. The phenomenon is further investigated in a numerical example in Section 4.3.

When no lower bound on c_s is used, the nonlocal equivalent strain artifacts just discussed also arise in the final stages of the inelastic regime, when damage accumulates and the stresses tend to zero, leading to a vanishing scaling factor s and scaled gradient parameter c_s. This is illustrated in Figure 23 that depicts the nonlocal equivalent strain fields of the final time steps indicated in the force-displacement curve in Figure 19 (in black). When c_s is limited to c_{min}, no
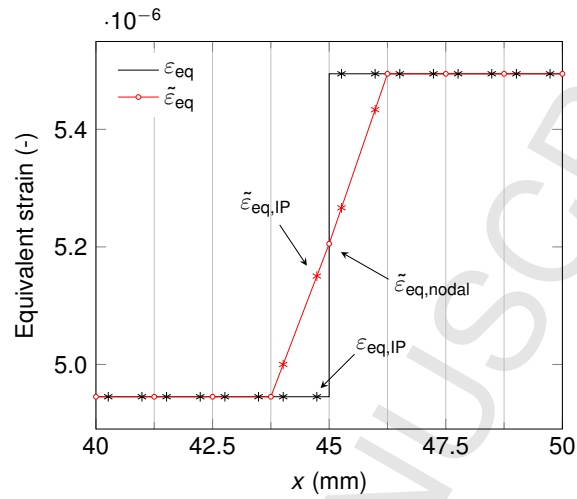
25

Figure 21: One-dimensional localization test (end-displacement $u = 0.0005$ mm): Local behavior of the nonlocal equivalent strain field when the gradient parameter equals c_{min} (close-up at the strain discontinuity on the left-hand side). The position of the integration points is marked by an asterisk. The vertical gray lines represent element boundaries.
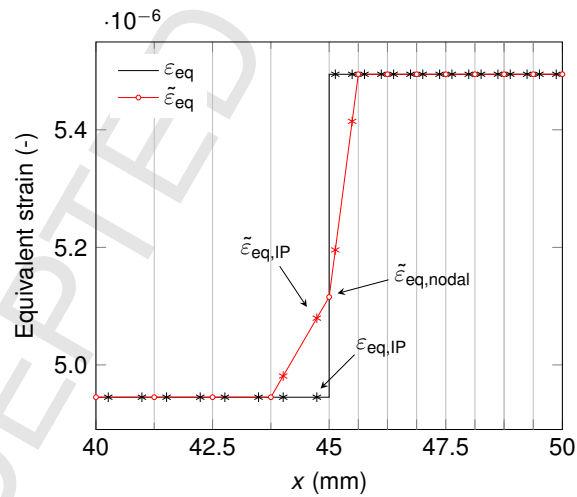

Figure 22: One-dimensional localization test using a non-uniform mesh (end-displacement $u = 0.0005$ mm): Local behavior of the nonlocal equivalent strain field when the gradient parameter equals c_{min} (close-up at the strain discontinuity on the left-hand side). The position of the integration points is marked by an asterisk. The vertical gray lines represent element boundaries.
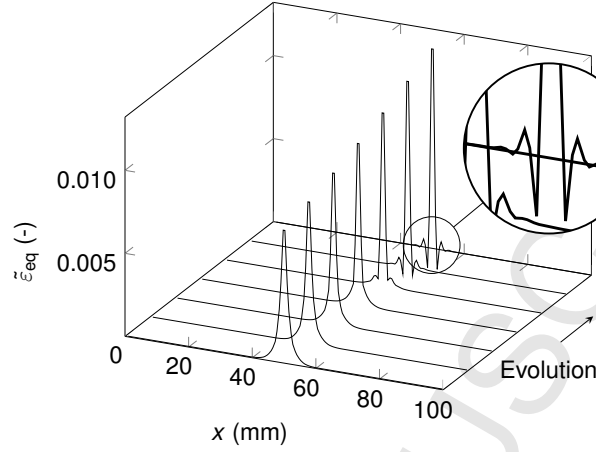
26

Figure 23: One-dimensional localization test: Nonlocal equivalent strain fields in the final stages of the simulation where c_s is not limited.

oscillations are present as can be observed from Figure 24, yet leading to a less brittle response as can be noticed from the red force-displacement curve in Figure 19.

Finally, it is worth noting that the use of a lumped system matrix to avoid oscillations in the solution is a common technique in other fields where transient and dynamic phenomena are modeled, such as water [22, 23] and heat [24, 25] flow. If a consistent non-diagonalized mass matrix is used for modeling these dynamic processes, positive off-diagonal terms and thus oscillations appear when the so-called discrete maximum principle is violated [26, 27]. This principle imposes a minimum time step size in the dynamic analysis that is apparently analogous to the use of a minimum length scale in the transient gradient-enhanced damage model.

4.2. Generalization to two dimensions

From the one-dimensional case in the previous section it was found that oscillations and nonlocal smoothing can be prevented by setting the off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$ to zero. In the same way, for two-dimensional analyses, exact minimum values of c_{xx}, c_{yy} and c_{xy} can be derived for three-node triangular (T3) elements since

$$\mathbf{K}_{\varepsilon\varepsilon} = \begin{bmatrix} k_{1,1} & & \text{sym.} \\ k_{2,1} & k_{2,2} & \\ k_{3,1} & k_{3,2} & k_{3,3} \end{bmatrix} \tag{46}$$

contains only three different off-diagonal terms. When a three-point Gauss integration quadrature rule is used and $\mathbf{c}_{\min}$ is kept equal for all the integration points, the following non-zero values for c_{xx}, c_{yy} and c_{xy} yield zero off-diagonal terms:

$$c_{xx,\min} = \frac{1}{6} \left(x_1{}^2 + x_2{}^2 + x_3{}^2 - x_1 x_2 - x_2 x_3 - x_3 x_1 \right) \tag{47}$$

$$c_{yy,\min} = \frac{1}{6} \left(y_1{}^2 + y_2{}^2 + y_3{}^2 - y_1 y_2 - y_2 y_3 - y_3 y_1 \right) \tag{48}$$
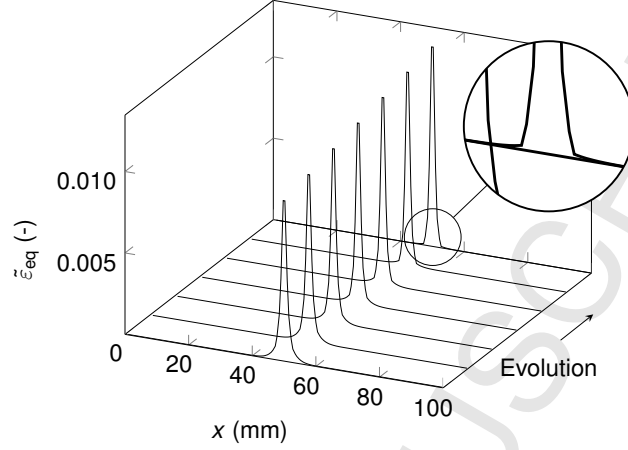
27

Figure 24: One-dimensional localization test: Nonlocal equivalent strain fields in the final stages of the simulation where c_s is limited to c_{min}.

and

$$c_{xy,min} = \frac{1}{6}(x_1y_1 + x_2y_2 + x_3y_3) - \frac{1}{12}(x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2) \quad (49)$$

in which (x_i, y_i) are the global coordinates of the finite element nodes. The meaning of these terms can be illustrated graphically: as mentioned in Section 3, these terms correspond to the shape and orientation of an ellipse with equation

$$(x - x_m)^2 \left(A_1 \cos^2 \theta_e + A_2 \sin^2 \theta_e\right) + (y - y_m)^2 \left(A_1 \sin^2 \theta_e + A_2 \cos^2 \theta_e\right)$$
$$+ 2(x - x_m)(y - y_m) \cos \theta_e \sin \theta_e (A_1 - A_2) = 1 \quad (50)$$

where the principal angle, which governs the orientation of the ellipse, is calculated according to

$$\theta_e = \frac{1}{2} \tan^{-1} \left(\frac{2c_{xy,min}}{c_{xx,min} - c_{yy,min}}\right), \quad (51)$$

x_m and y_m are the center of the ellipse, and the two factors

$$A_1 = \frac{1}{l_{1,min}^2} \qquad \text{and} \qquad A_2 = \frac{1}{l_{2,min}^2} \quad (52)$$

are defined by the principal axes lengths $l_{1,min} = \sqrt{2c_{1,min}}$ and $l_{2,min} = \sqrt{2c_{2,min}}$, where $c_{1,min}$ and $c_{2,min}$ are the eigenvalues of $\mathbf{c}_{min}$. Figure 25 provides a graphical interpretation of this minimum interaction ellipse $\mathbf{c}_{min}$ for various triangular finite element geometries, where the ellipses are centered at the finite element centroids. If the ellipse is centered at the integration points, as in Figure 26, all three ellipses pass through the other integration points of the element; this is analogous to what was observed in the one-dimensional case in the previous section (l_{min} was exactly the distance between the integration points).
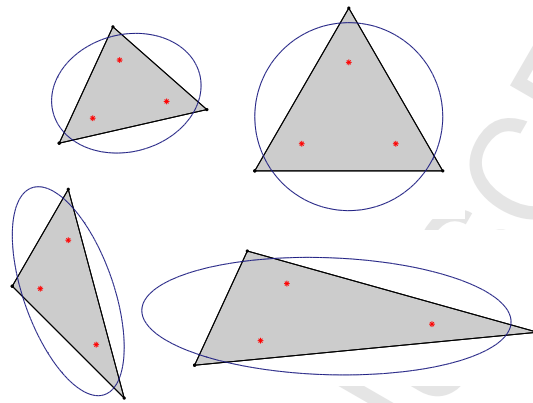
28

Figure 25: Graphical representation of $\mathbf{c}_{\min}$ for various finite element shapes. The ellipses are centered at the finite element centroids.
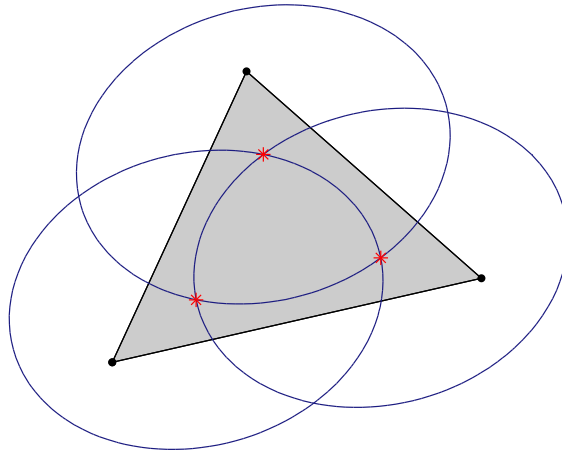
Figure 26: Graphical representation of $\mathbf{c}_{\min}$, where the ellipses are centered at the three Gauss integration points.

29

Although the value of $\mathbf{c}_{\min}$ given by Equations (47)-(49) will result in zero off-diagonal entries in the element stiffness matrix $\mathbf{K}_{\varepsilon\varepsilon}$, it will not lead to vanishing off-diagonal entries in the individual integration point stiffness matrices $\mathbf{K}_{\varepsilon\varepsilon,\text{IP1}}$, $\mathbf{K}_{\varepsilon\varepsilon,\text{IP2}}$, and $\mathbf{K}_{\varepsilon\varepsilon,\text{IP3}}$ as only the weighted sum of the off-diagonal terms will be zero. Appendix C reports the expressions of $\mathbf{c}_{\min,\text{IP}}$ that satisfy the vanishing off-diagonal term condition for each integration point individually.

In general, the off-diagonal terms do not become positive simultaneously during an analysis, leading to the question *when* an integration point (or finite element) should switch to a 'local' $\mathbf{c}_{\min}$. Numerical experience indicated that if $\mathbf{c}$ is set to $\mathbf{c}_{\min}$ when only one off-diagonal term becomes positive, nonlocal interaction is canceled too soon, leading to classical issues encountered with local damage models such as mesh-dependency of strain localization. On the other hand, if $\mathbf{c}$ is set to $\mathbf{c}_{\min}$ when only all three unique off-diagonal terms become positive, oscillations in the solution fields still take place. Therefore, a modified $\mathbf{c}'_{\min}$ is constructed that only turns positive off-diagonal terms to 0 (or a small negative residual ρ as explained later), while the other off-diagonal keep their original, negative value. More specifically, when evaluating the signs of the off-diagonal terms $k_{i,j}$ of an integration point stiffness matrix, which has the format expressed by (46), it can be observed that eight combinations exist for the sign of the three off-diagonal terms $k_{i,j}$ (all $k_{i,j}$ negative, all $k_{i,j}$ positive, $k_{2,1}$ positive whereas $k_{3,1}$ and $k_{3,2}$ negative, etc.), leading to eight unique corresponding values of $\mathbf{c}'_{\min}$. For example, when during an analysis a certain value of $\mathbf{c}$ (calculated according to (19) or (30)) leads to a positive $k_{3,2}$ and a negative $k_{2,1}$ and $k_{3,1}$, unique values for $c'_{xx,\min}$, $c'_{yy,\min}$ and $c'_{xy,\min}$ can be found that ensure $k'_{3,2} = 0$, $k'_{2,1} = k_{2,1}$ and $k'_{3,2} = k_{3,2}$, where a prime symbol indicates the modified or 'filtered' value. The inequality

$$k_{i,j} = p_{i,j}c_{xx} + q_{i,j}c_{yy} + r_{i,j}c_{xy} + s_{i,j} \leq \rho \, , \tag{53}$$

is checked for each off-diagonal term in every integration point and during every Newton-Raphson iteration. The factors are listed in Appendix D, and the residual ρ allows for a minimal nonlocal interaction between the nonlocal equivalent strain degrees of freedom at nodes i and j. Setting $\rho = 0$ will completely cancel nonlocal interaction but may lead to tortuous loading behavior in case of the principal stress-based model, as demonstrated in the example below. Indeed, even though oscillations in the nonlocal equivalent strain field are avoided by keeping the off-diagonal terms negative or zero, it has been observed that elements go into 'local mode' too early, leading to bifurcation behavior when damage grows, manifested by a saw-tooth force-displacement curve.

When $\rho = 0$ and all three unique off-diagonal terms are positive, $\mathbf{c}'_{\min}$ will be equal to $\mathbf{c}_{\min}$, therefore setting all three terms to zero. It can be understood that this specific value of $\mathbf{c}_{\min}$ effectively result in a fully lumped matrix $\mathbf{K}_{\varepsilon\varepsilon}$ where no coupling terms and consequently no interaction exists between the nonlocal equivalent strain degrees of freedom. In the case of, e.g., four-node quadrilateral elements, no combination of $c_{xx,\min}$, $c_{yy,\min}$ and $c_{xy,\min}$ exists that cancels all six unique off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$. Therefore, a diagonalized matrix can directly be computed using, for instance, a diagonally compensated reduction of positive off-diagonal entries [28]. When a positive off-diagonal term is detected in $\mathbf{K}_{ee,\text{IP}}$, this term is collapsed, i.e., set to zero and added to the corresponding diagonal entry. Unfortunately, this 'direct' manipulation of the stiffness matrix does not allow for a consistent linearization that is necessary to maintain quadratic convergence of the Newton-Raphson solution algorithm. However, the procedure described above leads to the filtered matrix

$$\mathbf{K}'_{\varepsilon\varepsilon} = \int_{\Omega} \left(\mathbf{N}_{\varepsilon}^{\text{T}}\mathbf{N}_{\varepsilon} + \mathbf{B}_{\varepsilon}^{\text{T}}\mathbf{c}'_{\min}\mathbf{B}_{\varepsilon} \right) d\Omega \tag{54}$$
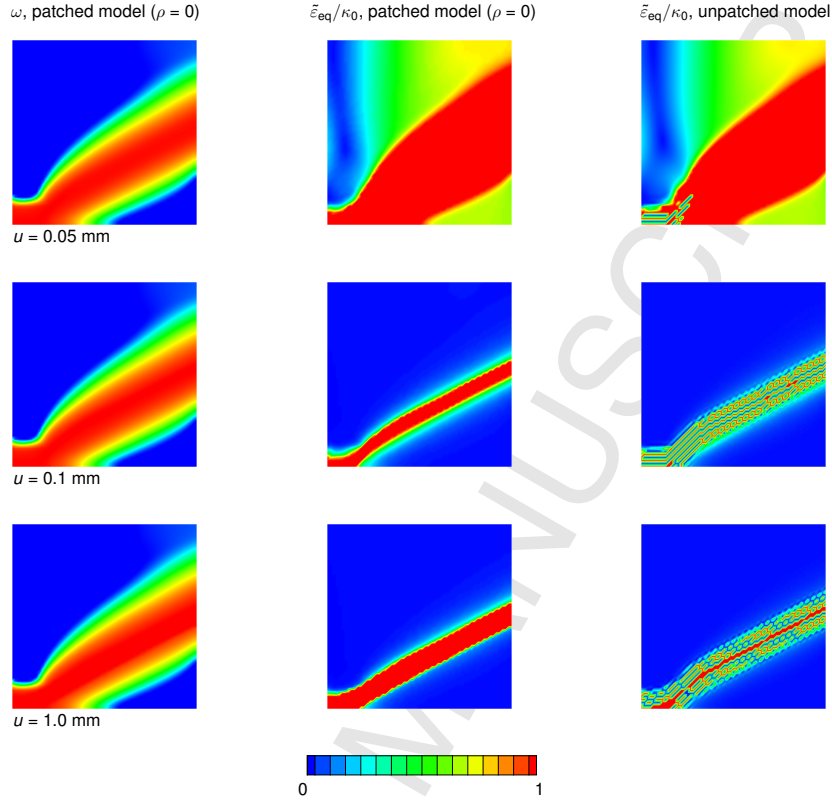
30

Figure 27: Shear band test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (central column) for the patched principal stress-based model with $\rho = 0$ (7200 elements). When comparing the normalized nonlocal equivalent strain fields with those from the unpatched model (right column), all field oscillations are canceled.

for T3 elements; the matrix contains no positive off-diagonal terms and can be easily linearized since an exact value of $\mathbf{c}'_{min}$ and its dependency on the solution fields $\mathbf{u}$ and $\tilde{\varepsilon}_{eq}$ exists, allowing for the evaluation of the derivative terms $\frac{\partial \mathbf{c}'_{min}}{\partial \mathbf{u}}$ and $\frac{\partial \mathbf{c}'_{min}}{\partial \tilde{\varepsilon}_{eq}}$ (see Appendix B).

4.3. Model performance: Shear band problem

The patched models, developed in the previous section, are again applied to the shear band problem introduced in Section 3.3. The same meshes and material parameters are employed. When using the patched stress-based model with residual parameter $\rho = 0$ (thus canceling all off-diagonal terms when the length scale becomes too small), it can be seen from the central column of the snapshots in Figure 27 that the nonlocal strain field does not show any oscillations when compared with the unpatched version of the model (right column). Oscillations are thus canceled, while still reducing the spreading of the damage band, as can be noticed from the damage fields in the left column of the figure. The force-displacement curves in Figure 28, however, still suffers from a tortuous character when using finer meshes, indicating that the solution fields lack sufficient smoothness. It appears that, using a principal stress-based definition of the transient
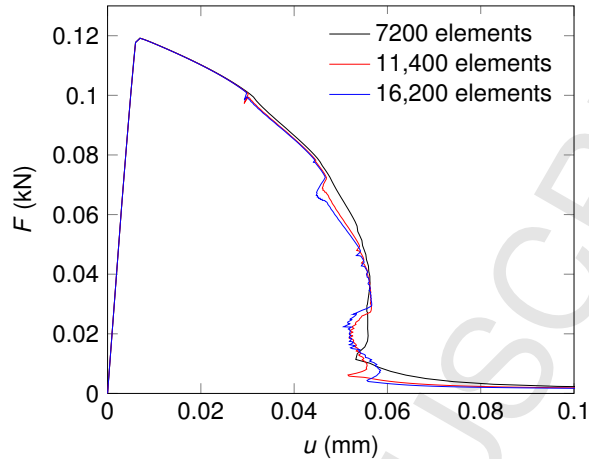
31

Figure 28: Shear band test using the patched principal stress-based model with $\rho = 0$: Force-displacement curves for various mesh sizes.

gradient parameter (Equation (19)), nonlocal interaction is canceled too early, causing typical bifurcation phenomena where the algorithm has difficulties in choosing in which finite element strain localization should occur. This tortuous behavior can be mitigated either by using a non-zero residual interaction parameter ρ, allowing for residual nonlocal interaction, or by resorting to the nonlocal equivalent stress-based model, where the transient gradient parameter is more at pace with the damage-driving parameter. Indeed, for shear-dominant failure, the principal-stress based model cancels nonlocal interaction too soon with respect to the damage-driving parameter $\tilde{\varepsilon}_{eq}$ that depends on the definition of the equivalent strain. This can be demonstrated when comparing the evolutions of the damage, expressed by (6), and the gradient parameters, defined by Equations (27)-(29). Figure 29 indicates that for a pure shear state ($\varepsilon_{xx} = \varepsilon_{yy} = 0$, $\varepsilon_{xy} \neq 0$) and using the material parameters $k = 1$, $\kappa_0 = 0.0001$, and $\nu = 0$, the diffusive behavior, reflected by the ratio $c_{xx}/c = c_{yy}/c$ (c_{xy} equals zero), is much smaller than in case of a pure tensile state ($\varepsilon_{xy} = \varepsilon_{yy} = 0$, $\varepsilon_{xx} \neq 0$). On the other hand, the nonlocal equivalent stress-based model will always lead to the same scaling parameter $s = c_{xx}/c = c_{yy}/c$ for a given damage value, regardless of the strain (or stress) state. This observation speaks in favour of using the nonlocal equivalent stress-based model, since it goes hand-in-hand with the employed definition of the equivalent strain. Figure 30 shows that the patched version of this model does not show any oscillations when compared with the unpatched version, and significant damage spreading does not take place at higher deformation levels (top displacement of 1.0 mm). In contrast with the principal stress-based model, the force-displacement curve (Figure 31) has no tortuous character and shows proper convergence upon mesh refinement. When using the principal stress-based model with residual interactions, a fully smooth force-displacement curve is obtained when the residual parameter $\rho = -0.1$, as depicted in Figure 32, but at the expense of more damage spreading at the final stages of the simulation (Figure 32). Intermediate values of ρ still lead to some degree of tortuous response, as indicated by Figures 34 and 35.

Next, the analysis is performed for the special case when $\mathbf{c}$ is fixed to $\mathbf{c}_{min}$ during the analysis. As expected, the model behaves in a local manner and the force-displacement curves in Figure 36 do not converge upon mesh refinement because less energy is dissipated when using
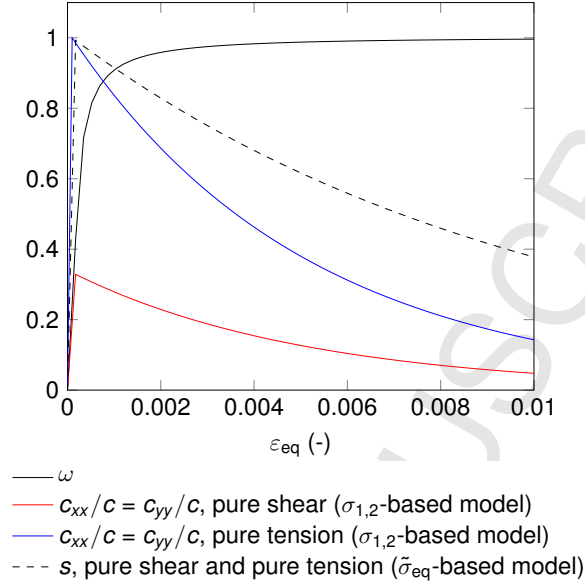
32

Figure 29: Evolution of the damage parameter ω and the amount of nonlocal interaction for pure shear and pure tensile states, represented by the ratios $c_{xx}/c = c_{yy}/c$ for the principal stress-based model and the parameter s for the nonlocal equivalent stress-based model.

finer meshes; this is also reflected in the damage profiles in Figure 37: damage remains contained in a narrower shear band for the finest discretizations. Remarkably, setting $\mathbf{c} = \mathbf{c}_{\min}$ still leads to spreading of the damage band at high deformation levels, as can be noticed from the bottom row of damage fields in Figure 37. Since no direct interaction is possible between nonlocal strains due to the absence of coupling terms in $\mathbf{K}_{\varepsilon\varepsilon}$, the widening of the damage band is purely caused by the (piecewise) continuous character of the nonlocal equivalent strain field, as discussed in Section 4.1: elements exhibiting high strain and thus damage levels can 'infect' neighboring elements because the damage-driving, nonlocal equivalent strain levels are averaged between two adjacent elements. The damage spreading phenomenon can thus not fully be resolved by setting $\mathbf{c}$ to $\mathbf{c}_{\min}$: the inherent continuous character of the nonlocal equivalent strain field will eventually result in spreading, albeit less significant.

Finally, when using the nonlocal equivalent stress-based model in conjunction with a dedicated evolution law for the nonlocal interaction behavior (38), a very thin damage band can be achieved, analogous to the transient model by Poh and Sun [8], as was observed in Figures 11-13. Indeed, by using a very high β_s parameter in (38), which differs from the β parameter of the damage evolution law (6), nonlocal interaction is canceled very rapidly after damage initiates, as is also the case when using the micromorphics-based scaling parameter expressed by (15). Figure 40 illustrates this behavior for $\beta_s = 10,000$ and $c = 9$ mm^2, using the finest discretization (16,200 elements): the width of the total damage band is comparable to the stress-based models without a dedicated evolution law for the nonlocal interaction but the region of highly damaged material is much more concentrated along the shear band. Again, at very large deformation levels ($u = 1$ mm, right snapshot in Figure 40), this narrow damage band tends to spread. Moreover, because the degree of nonlocal interaction is reduced very quickly, the patched model switches
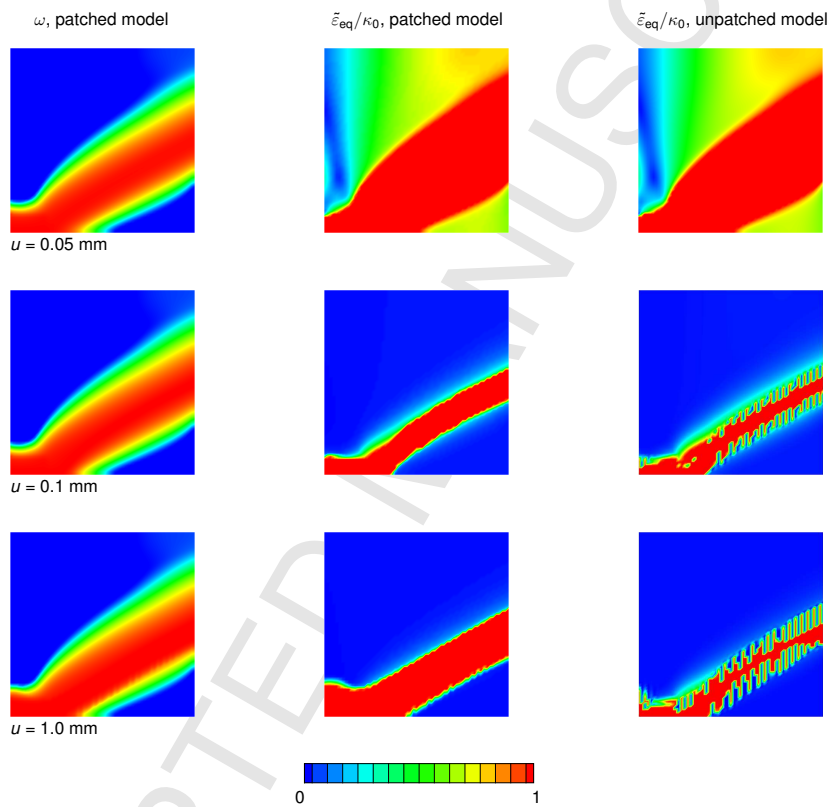
33

Figure 30: Shear band test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (central column) for the patched nonlocal equivalent stress-based model with $\rho = 0$ (7200 elements). When comparing the normalized nonlocal equivalent strain fields with those from the unpatched model (right column), all field oscillations are canceled.
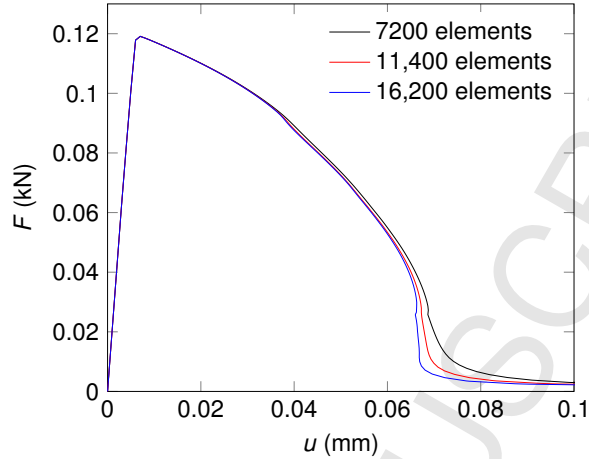
Figure 31: Shear band test using the patched nonlocal equivalent stress-based model with $\rho = 0$: force-displacement curves for various mesh sizes.

to its 'local mode' ($\mathbf{c} = \mathbf{c}_{min}$) in an early stage, resulting in no proper convergence upon mesh refinement of the force-displacement curves in Figure 38, comparable to what was observed when setting $\mathbf{c} = \mathbf{c}_{min}$ throughout the entire simulation. In case of the unpatched model, the convergence behavior is improved, as can be noticed from the loading behavior in Figure 39.

5. Further analyses of damage initiation and propagation

5.1. Analysis of damage initiation: Tensile test

In this example, the issue of wrong damage initiation reported by Simone et al. [5] is reinvestigated using the stress-based models. The test setup, depicted by Figure 41, consists of a notched rectangular specimen under tensile loading (elongation of 0.002 mm). The domain is meshed using 6282 triangular elements with Young's modulus $E = 1000$ N/mm^2 and Poisson's ratio $\nu = 0$. The gradient parameter $c = 0.02$ mm^2 and the modified von Mises material model (3) is used with $k = 10$ and $\kappa_0 = 0.2$. As shown in Figure 42, the standard gradient-enhanced damage model gives rise to a profile of the nonlocal equivalent strain that does not attain its maximum at the notch tip even if the local equivalent strain is there unbounded according to linear elastic fracture mechanics (a similar result [5] is obtained with the integro-differential nonlocal damage model [1]). Since experimental evidence indicates that in notched specimens made of a quasi-brittle material a crack initiates at the notch tip [29], this observation about the location of the maximum of the field driving damage initiation implies that classical nonlocal models, in either integro-differential or differential form, lead to physically inconsistent results. In contrast to the standard gradient-enhanced model, for both stress-based models the peak of the nonlocal equivalent strain is located at the crack tip, leading to an initiation of damage at the correct location. This example also shows the importance of using a lower bound on $\mathbf{c}$: when no lower bound is used, oscillations and even unphysical negative values arise in the profile of the nonlocal equivalent strain as shown in Figure 43.
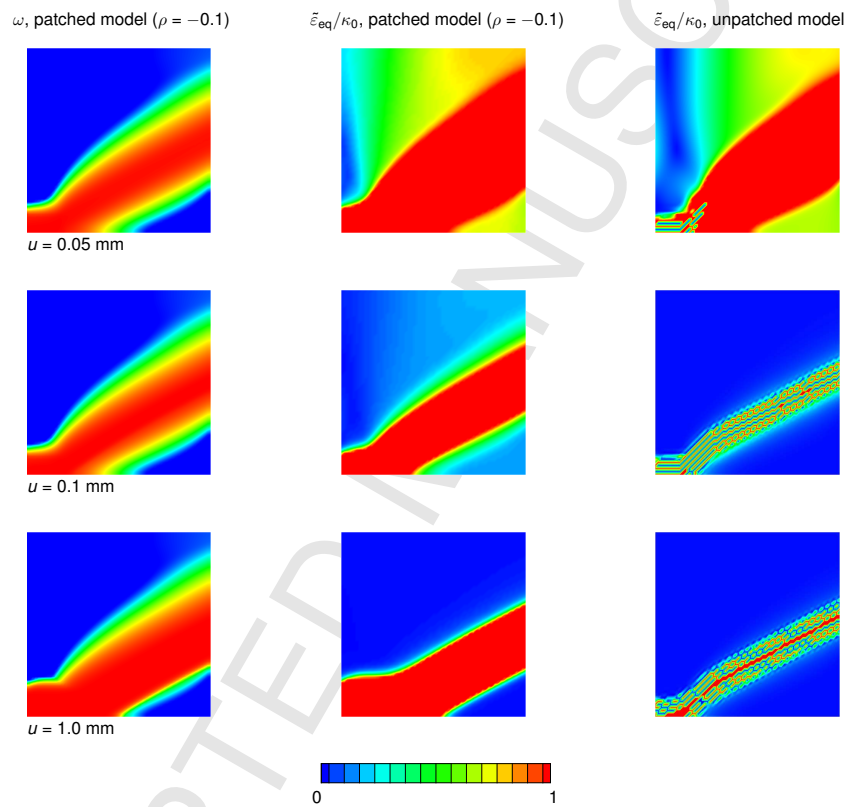
35

Figure 32: Shear band test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (central column) for the patched principal stress-based model with $\rho = -0.1$ (7200 elements). When comparing the normalized nonlocal equivalent strain fields with those from the unpatched model (right column), all field oscillations are canceled.
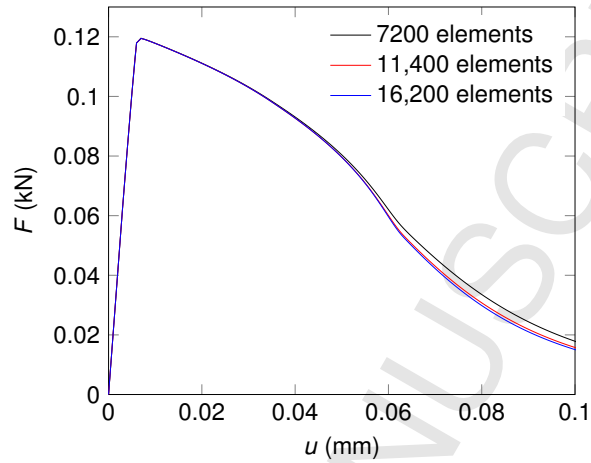
Figure 33: Shear band test using the patched principal stress-based model with $\rho = -0.1$: Force-displacement curves for various mesh sizes.
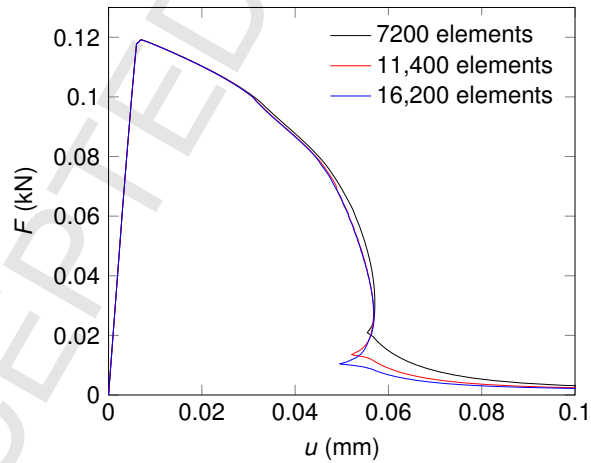

Figure 34: Shear band test using the patched principal stress-based model with $\rho = -0.01$: Force-displacement curves for various mesh sizes.
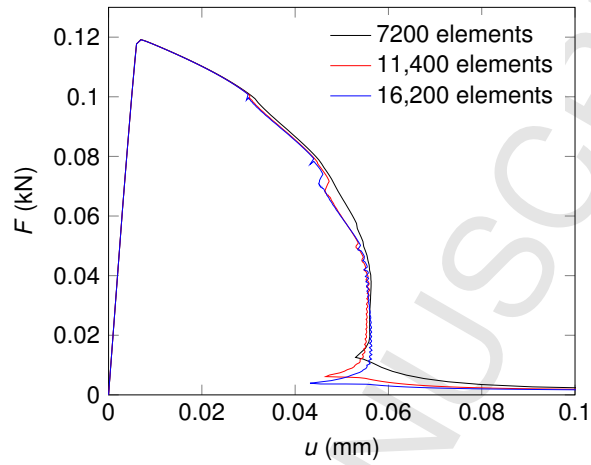
Figure 35: Shear band test using the patched principal stress-based model with $\rho = -0.001$: Force-displacement curves for various mesh sizes.
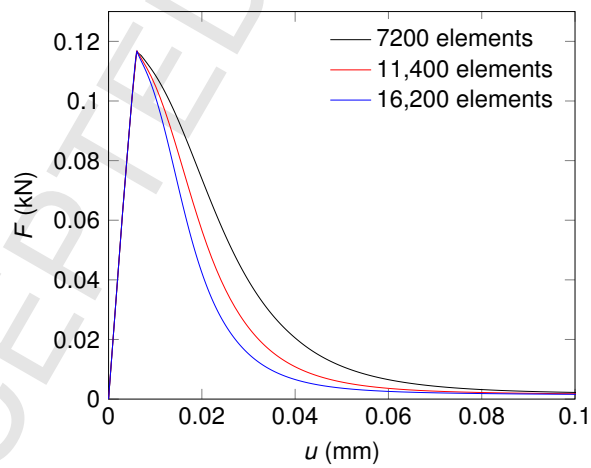

Figure 36: Shear band test using $\mathbf{c} = \mathbf{c}_{\min}$ in all elements: Force-displacement curves for various mesh sizes.

38

7200 elements   11,400 elements   16,200 elements
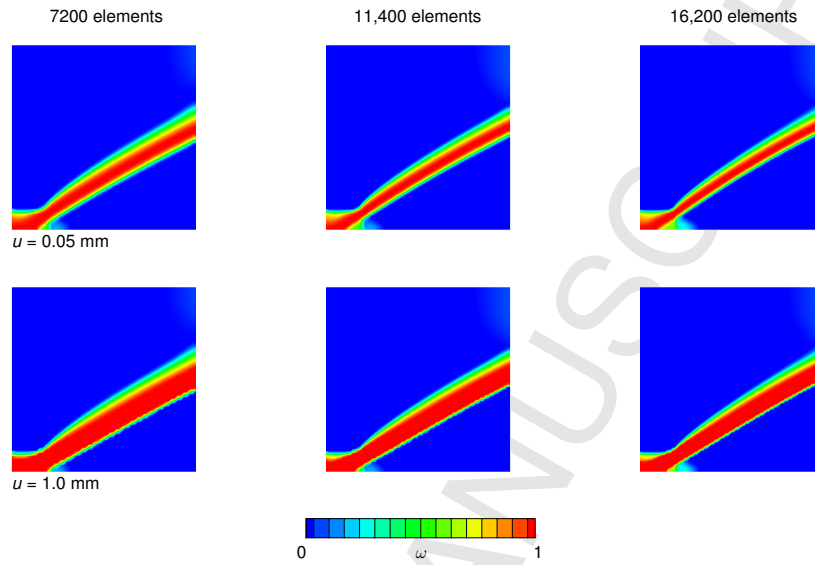
u = 0.05 mm

u = 1.0 mm

0   ω 1

Figure 37: Shear band test using $\mathbf{c} = \mathbf{c}_{\min}$ in all elements: Damage fields at imposed displacement levels (0.05 mm and 1.0 mm), for meshes containing 7200 (left column), 11,400 (central column) and 16,200 elements (right column).
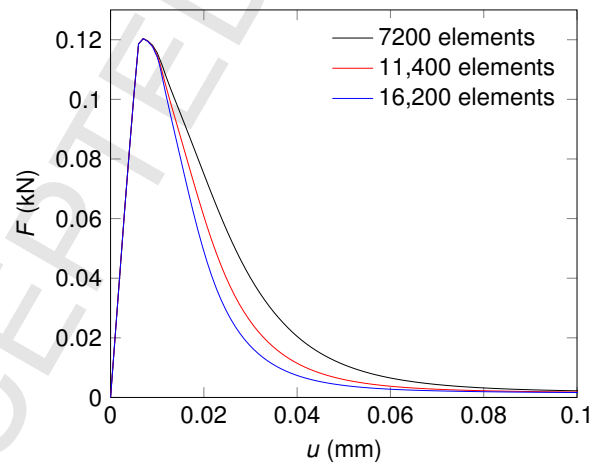

Figure 38: Shear band test using the patched nonlocal equivalent stress-based model with a dedicated evolution law for the diffusivity activity: Force-displacement curves for various mesh sizes.

39

Figure 39: Shear band test using the unpatched nonlocal equivalent stress-based model with a dedicated evolution law for the diffusivity activity: Force-displacement curves for various mesh sizes.
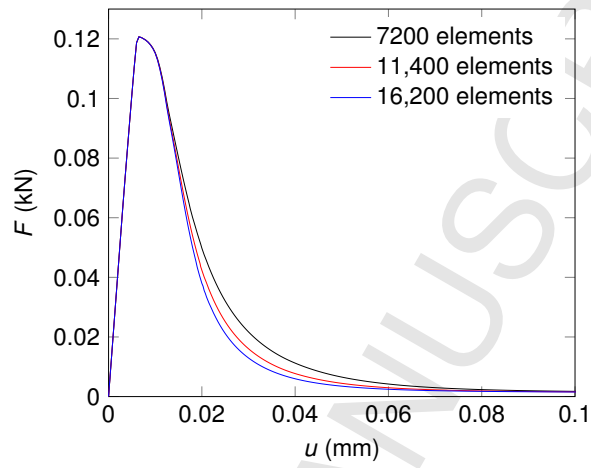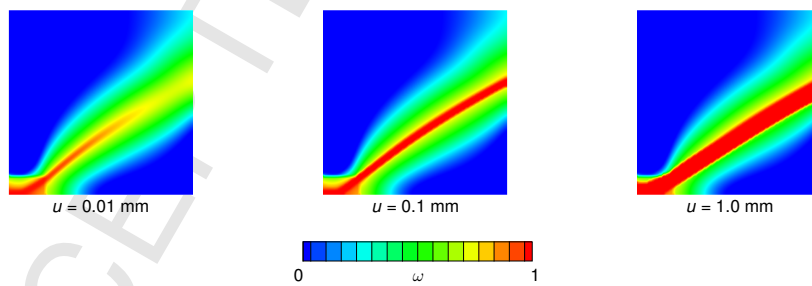

Figure 40: Shear band test using the patched nonlocal equivalent stress-based model with a dedicated evolution law for the diffusivity activity: Evolution of the damage field.
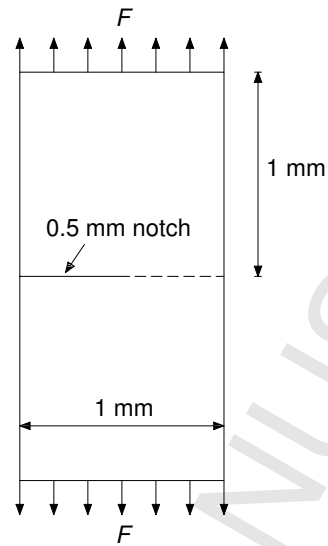
40

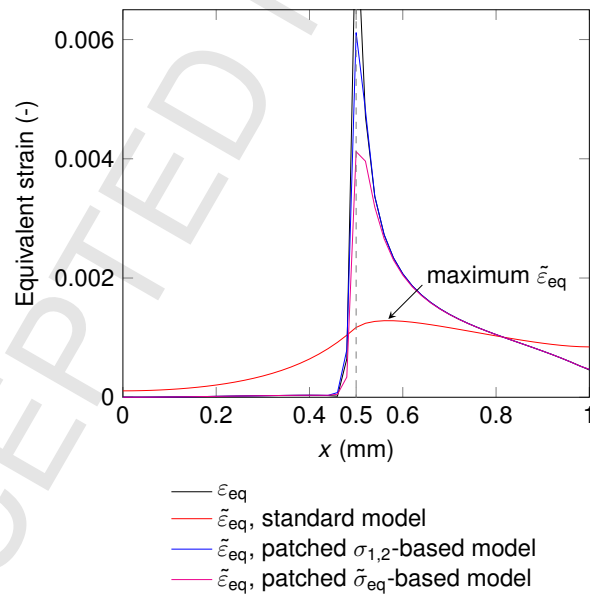Figure 41: Compact tension test: Geometry and boundary conditions.

Figure 42: Compact tension test: Local and nonlocal equivalent strain fields along the horizontal symmetry line of the specimen.
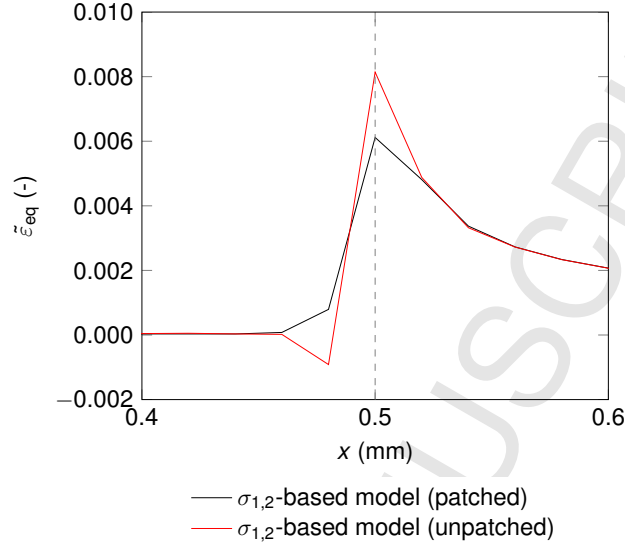
41

Figure 43: Compact tension test using the unpatched and patched principal stress-based model: Nonlocal equivalent strain fields along the horizontal symmetry line of the specimen. Note that the displayed spatial range is limited to an interval of 0.2 mm.
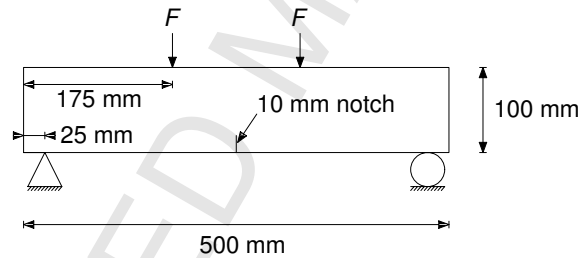
Figure 44: Four-point bending test: Geometry and boundary conditions.

5.2. Analysis of damage propagation: Four-point bending test

As a final example, the stress-based models are used to model a concrete beam under four-point bending. As depicted in Figure 44, damage is initiated at the midspan of the beam using a 10 mm notch. Plane stress conditions are assumed with Young's modulus $E = 40,000$ N/mm^2 and Poisson's ratio $\nu = 0.2$. A gradient activity parameter $c = 4$ mm^2 has been employed and the nonlinear material behavior is governed by the modified von Mises damage model presented in Section 2.1 ($\kappa_0 = 0.000075$, $k = 10$, $\alpha = 0.92$, and $\beta = 300$). The model is discretized using 39,202 triangular elements in which the displacements and nonlocal equivalent strains are interpolated using quadratic and linear shape functions, respectively. When applying the standard model to this problem, unphysical damage spreading takes place near the bottom of the specimen, as can be noticed from Figure 45.

The resulting force-displacement curves in Figure 46 confirm the more brittle behavior of the proposed models already observed in the previous examples. The curves also show that the load-

42

Figure 45: Four-point bending test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column) for the standard model.
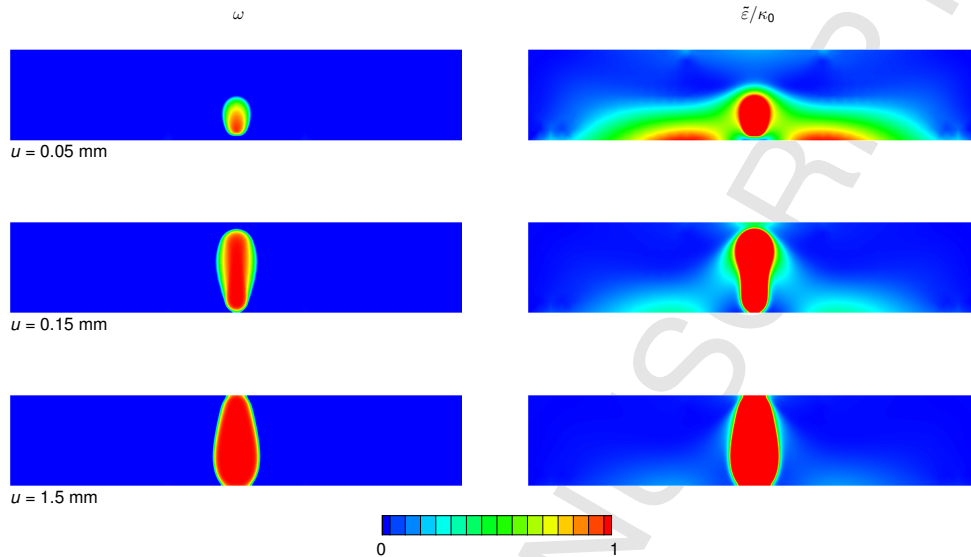
ing behavior of the patched stress-based models does not differ significantly from the unpatched versions, although the patched versions effectively cancel the nonlocal equivalent strain field oscillations present in the unpatched versions. In case of the principal stress-based model, for instance, the nonlocal strain fields in Figure 47 suffer again from oscillations near the localization band, whereas these oscillations are not present in case of the unpatched model (Figure 48). In contrast with the shear band problem in Section 4.3, no residual interaction parameter ρ is necessary to obtain smooth force-displacement curves; this evidences the conclusion that the principal stress-based model works better in case of mode-I failure problems, i.e., the transition to local behavior does not happen prematurely with respect to damage, contrary to the case of mode-II-dominant failure highlighted in Figure 29. From the evolution of the damage fields in Figures 47 and 48, it can be concluded that the unpatched principal stress-based model does not lead to any significant spreading of damage thanks to the aforementioned 'diking effect' of the nonlocal strain field oscillations. On the other hand, the oscillation-free model does show some minor spreading at high values of the imposed top displacement (bottom left snapshot in Figure 48).

Finally, in case of the nonlocal equivalent stress-based model, the dedicated evolution law (38) with $\beta_s = \beta$ is used for the scaling factor s since the relatively low value of α (= 0.92) would result in residual nonlocal interactions and thus increase the amount of unphysical damage spreading, as discussed in Section 3.2. As can be noted from the evolution of s in the right column of the snapshots in Figure 49, nonlocal interaction (or diffusivity) takes the form of an upward moving crescent and gets close to one near the crack tip while evolves to zero along a band of localized strains. Unfortunately, s grows near the bottom of the specimen too, although this does not lead to any significant spreading of damage as can be noticed from the bottom left snapshot. When compared with the loading behavior of a model that uses the same evolution law for damage and nonlocal interaction, the force-displacement curves in Figure 50 indicate that there is little
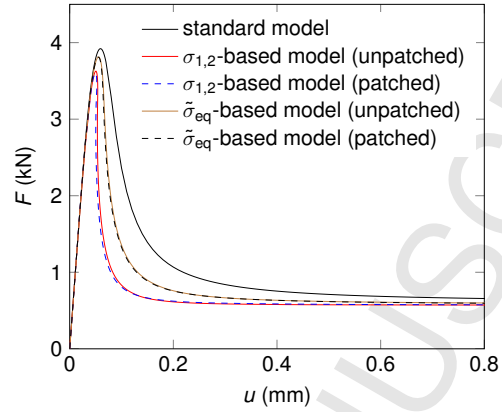
43

Figure 46: Four-point bending test: Force-displacement curves.

Figure 47: Four-point bending test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column) for the unpatched principal stress-based model.

44

ω $\bar{\varepsilon}/\kappa_0$

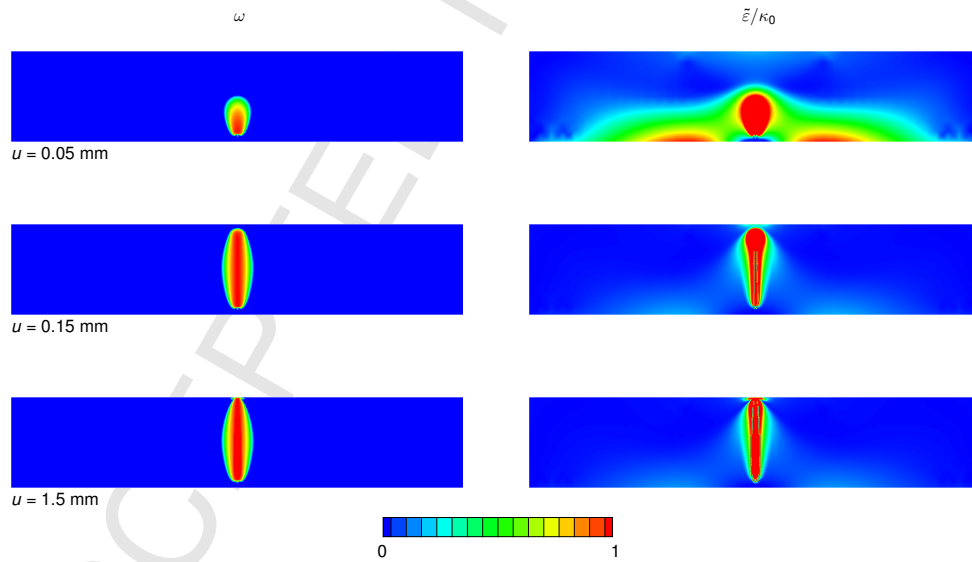$u = 0.05$ mm
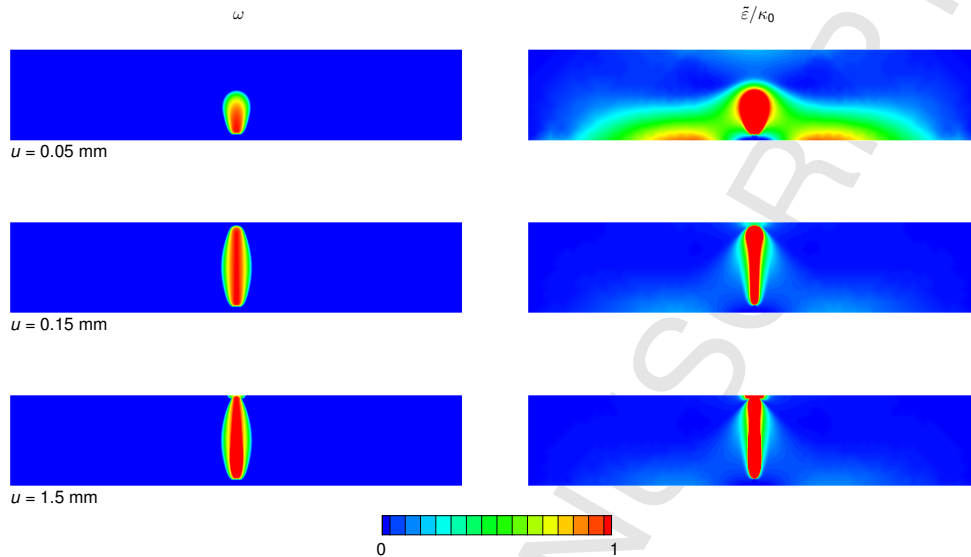
$u = 0.15$ mm

$u = 1.5$ mm

0   1

Figure 48: Four-point bending test: Evolution of the damage (left column) and normalized nonlocal equivalent strain fields (right column) for the patched principal stress-based model.

difference between both models.

6. Summary and concluding remarks

The objective representation of strain localization is an essentiality when modeling failure in quasi-brittle solids. The gradient-enhanced damage model by Peerlings et al. [2] has excellent regularization properties when modeling strain localization but suffers from some issues such as unphysical spreading of damage and wrong location of damage initiation. In this paper, two stress-based transient gradient-enhanced damage models were introduced to alleviate these problems: a model based on the principal stresses, which was directly inspired by the stress-based integral nonlocal model by Giry et al. [12], and a nonlocal equivalent stress-based model that uses a scaling factor related to the adopted nonlocal equivalent strain measure. Both anisotropic models rely on the reduction of the gradient parameter and thus the amount of nonlocality when stresses decrease.

It was shown that a vanishing gradient parameter results in oscillations in the damage-driving, nonlocal equivalent strain field. Although these artifacts may help in avoiding spreading of damage, they are unphysical and mesh size-dependent. Therefore, patched versions of the discretized stress-based models are developed to introduce a finite element size-dependent minimum value of the gradient parameter c or gradient matrix **c**. It has been demonstrated that the patched principal stress-based model does not perform well in case of shear-dominant failure and leads to tortuous force-displacement curves since the degree of nonlocal interaction is too low with respect to the damage level. A principal stress-based model with residual interactions has been proposed to mitigate this phenomenon, yet it reintroduces the damage spreading issue. The nonlocal equivalent stress-based model performs better since the evolution of diffusivity is linked to the measure of the local equivalent strain that, through the diffusion equation, governs the

45

ω s

$u = 0.05$ mm

$u = 0.15$ mm

$u = 1.5$ mm

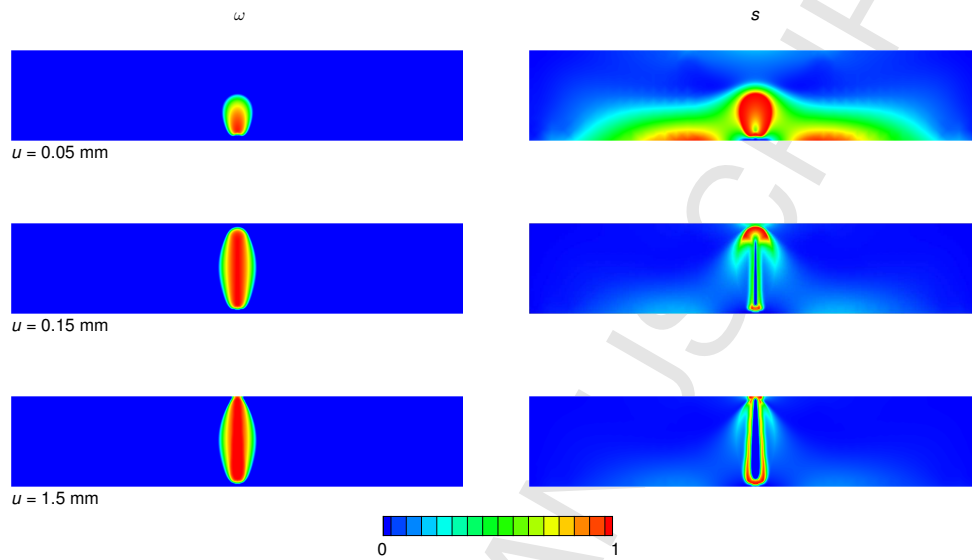0   1

Figure 49: Four-point bending test: Evolution of the damage (left column) and scaling parameter fields (right column) for the patched nonlocal equivalent stress-based model with dedicated nonlocal interaction law.
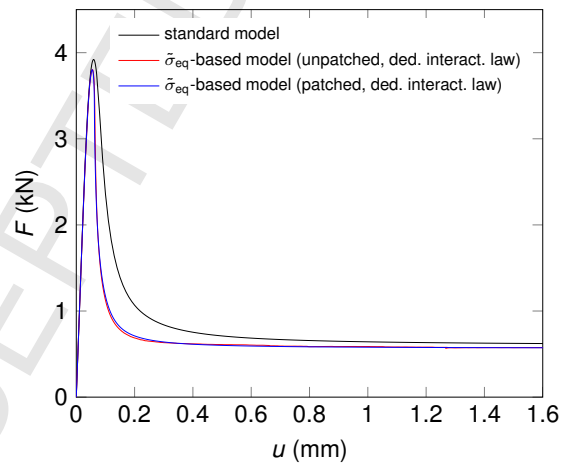

Figure 50: Four-point bending test: Force-displacement curves for the nonlocal equivalent stress-based models with and without dedicated nonlocal interaction law.

46

damage. The nonlocal equivalent stress-based model, in its general form, can be considered as a generic anisotropic gradient-enhanced model that can be equipped with a dedicated evolution law for the nonlocal interaction scaling factor. This is necessary to completely cancel nonlocal interactions when a damage evolution law that allows for residual stresses is used.

Through the simulation of a notched rectangular specimen loaded under tension, it was demonstrated that the patched stress-based models lead to a correct location of damage initiation, in contrast with the standard gradient-enhanced damage model where damage initiates some distance away from the notch tip. The patched models also reduce the amount of unphysical damage spreading significantly, although spreading still takes place at high loading levels. It was shown that this spreading is not caused by residual diffusion but by the inherent piecewise continuous character of the damage-driving nonlocal equivalent strain field: a discrete jump in the local equivalent strain field over an element border will lead to an averaged value of the nonlocal equivalent strain field at that border, even when nodal nonlocal interaction and therefore diffusion spreading is completely canceled by using a lumped system matrix. This continuity spreading phenomenon is consequently caused by the employed discretization technique and is not an inherent feature of the proposed models.

Acknowledgement

Appendix A. Principal stress-based model: Simplified expressions for c_{xx}, c_{yy}, and c_{xy}

For the principal stress-based gradient-enhanced damage model, the components of $\mathbf{c}_\sigma$,

$$c_{xx} = c_1 \cos^2 \theta_p + c_2 \sin^2 \theta_p \,, \tag{A.1}$$

$$c_{yy} = c_1 \sin^2 \theta_p + c_2 \cos^2 \theta_p \,, \tag{A.2}$$

and

$$c_{xy} = c_{yx} = (c_1 - c_2) \cos \theta_p \sin \theta_p \,, \tag{A.3}$$

can be simplified by considering the trigonometric relations

$$\cos^2\left(\frac{\alpha}{2}\right) = \frac{1 + \cos \alpha}{2} \,, \tag{A.4}$$

$$\sin^2\left(\frac{\alpha}{2}\right) = \frac{1 - \cos \alpha}{2} \,, \tag{A.5}$$

$$\cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\alpha}{2}\right) = \frac{\sin \alpha}{2} \,, \tag{A.6}$$

$$\cos\left(\tan^{-1} x\right) = \frac{1}{\sqrt{1 + x^2}} \,, \tag{A.7}$$

and

$$\sin\left(\tan^{-1} x\right) = \frac{x}{\sqrt{1 + x^2}} \,. \tag{A.8}$$

Combining Equations (A.4)-(A.6) and (A.7)-(A.8) and setting $\alpha = \tan^{-1} x$ yield

$$\cos^2\left(\frac{\tan^{-1} x}{2}\right) = \frac{1 + \frac{1}{\sqrt{1+x^2}}}{2} \,, \tag{A.9}$$

$$\sin^2\left(\frac{\tan^{-1} x}{2}\right) = \frac{1 - \frac{1}{\sqrt{1+x^2}}}{2} \,, \tag{A.10}$$

and

$$\cos\left(\frac{\tan^{-1} x}{2}\right)\sin\left(\frac{\tan^{-1} x}{2}\right) = \frac{x}{2\sqrt{1 + x^2}} \,. \tag{A.11}$$

Replacing x in Equations (A.9)-(A.11) by $\frac{2\tau_{xy}}{\sigma_{xx}-\sigma_{yy}}$ and referring to the expression of the principal angle (23) result in

$$\cos^2 \theta_p = \frac{1}{2} + \frac{\sigma_{xx} - \sigma_{yy}}{4R} \,, \tag{A.12}$$

$$\sin^2 \theta_p = \frac{1}{2} - \frac{\sigma_{xx} - \sigma_{yy}}{4R} \,, \tag{A.13}$$

and

$$\cos \theta_p \sin \theta_p = \frac{\tau_{xy}}{2R} \,, \tag{A.14}$$

48

in which $R = \sqrt{\left(\frac{\sigma_{xx}-\sigma_{yy}}{2}\right)^2 + \tau_{xy}^2}$. Inserting Equations (21) and (A.12)-(A.14) into (A.1)-(A.2) yields, after some algebraic manipulations, the following simplified expressions for the components of $\mathbf{c}$:

$$c_{xx} = \frac{c}{f_t^2}\left(\sigma_{xx}^2 + \tau_{xy}^2\right), \tag{A.15}$$

$$c_{yy} = \frac{c}{f_t^2}\left(\sigma_{yy}^2 + \tau_{xy}^2\right), \tag{A.16}$$

and

$$c_{xy} = \frac{c}{f_t^2}\tau_{xy}\left(\sigma_{xx} + \sigma_{yy}\right). \tag{A.17}$$

Appendix B. Finite element discretization and consistent linearization of the stress-based models

Finite element discretization. The models presented in this work are discretized using a Galerkin finite element approach. To this end, the strong forms of the equilibrium equation without body forces (12) and diffusion equation (16) are converted in their respective weak forms,

$$\int_\Omega \mathbf{w}_u \cdot (\nabla \cdot \boldsymbol{\sigma})\, d\Omega = 0, \tag{B.1}$$

and

$$\int_\Omega w_\varepsilon \tilde{\varepsilon}_{eq} d\Omega - \int_\Omega w_\varepsilon \left(\nabla \cdot \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right)\right) d\Omega = \int_\Omega w_\varepsilon \varepsilon_{eq} d\Omega, \tag{B.2}$$

where $\mathbf{w}_u$ and w_ε are weight functions, in the domain Ω. Taking into account the identity

$$\mathbf{w}_u \cdot (\nabla \cdot \boldsymbol{\sigma}) = \nabla \cdot (\mathbf{w}_u\boldsymbol{\sigma}) - (\nabla\mathbf{w}_u) : \boldsymbol{\sigma} \tag{B.3}$$

and applying Green's formula, the weak form of the equilibrium equation (B.1) can be restated as

$$\int_\Omega (\nabla\mathbf{w}_u) : \boldsymbol{\sigma}\, d\Omega - \int_{\Gamma_t} \mathbf{w}_u \cdot \bar{\mathbf{t}}\, d\Gamma = 0 \tag{B.4}$$

where $\bar{\mathbf{t}} = \boldsymbol{\sigma} \cdot \mathbf{n}$ are the tractions acting on the domain boundary Γ_t oriented according to the outward unit normal vector $\mathbf{n}$. Using Green's formula again, the second term of the left-hand side in (B.2) can be reformulated as

$$\int_\Omega w_\varepsilon \left(\nabla \cdot \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right)\right) d\Omega = - \int_\Omega (\nabla w_\varepsilon) \cdot \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) d\Omega + \int_\Gamma w_\varepsilon \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) \cdot \mathbf{n} d\Gamma \tag{B.5}$$

where $\Gamma = \Gamma_u \cup \Gamma_t$ is the domain boundary composed of a boundary part with imposed displacements Γ_u and a boundary part with imposed tractions Γ_t. Combining Equations (B.2) and (B.5) results in the reworked weak form of the diffusion equation

$$\int_\Omega w_\varepsilon \tilde{\varepsilon}_{eq} d\Omega + \int_\Omega (\nabla w_\varepsilon) \cdot \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) = \int_\Omega w_\varepsilon \varepsilon_{eq} d\Omega + \int_\Gamma w_\varepsilon \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) \cdot \mathbf{n} d\Gamma. \tag{B.6}$$

In line with conventional isotropic gradient-enhanced damage models, the natural boundary condition $\left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) \cdot \mathbf{n} = 0$ is imposed on Γ so that the boundary integral cancels from the weak form of the field equation (B.6), resulting in

$$\int_\Omega w_\varepsilon \tilde{\varepsilon}_{eq} d\Omega + \int_\Omega (\nabla w_\varepsilon) \cdot \left(\mathbf{c}\nabla\tilde{\varepsilon}_{eq}\right) = \int_\Omega w_\varepsilon \varepsilon_{eq} d\Omega. \tag{B.7}$$

This boundary conditions implies that there is no flux of nonlocal equivalent strain through the boundary Γ.

The weak form of the equilibrium equation (B.4) can be discretized using finite elements by considering the interpolated value of the displacements

$$\mathbf{u} = \mathbf{N}_u \mathbf{u}^* \tag{B.8}$$

and the corresponding weight functions

$$\mathbf{w}_u = \mathbf{N}_u \mathbf{w}_u^* \tag{B.9}$$

where $\mathbf{N}_u$ is a shape function matrix that interpolates the nodal displacements $\mathbf{u}^*$ and $\mathbf{w}_u^*$. The resulting discretized weak form of the equilibrium equation reads, using matrix notation,

$$\int_\Omega (\mathbf{w}_u^*)^{\mathrm{T}} \mathbf{B}_u^{\mathrm{T}} \sigma \mathrm{d}\Omega - \int_\Gamma (\mathbf{w}_u^*)^{\mathrm{T}} \mathbf{N}_u^{\mathrm{T}} \mathbf{t} \mathrm{d}\Gamma = 0 \tag{B.10}$$

where $\mathbf{B}_u$ is a matrix containing the derivatives of the shape functions in $\mathbf{N}_u$. Since (B.10) must hold for any arbitrary weight function $\mathbf{w}_u^*$, the discretized equilibrium equation becomes

$$\int_\Omega \mathbf{B}_u^{\mathrm{T}} \sigma \mathrm{d}\Omega - \int_\Gamma \mathbf{N}_u^{\mathrm{T}} \mathbf{t} \mathrm{d}\Gamma = \mathbf{0} . \tag{B.11}$$

The weak form of the diffusion equation is discretized analogously using the interpolated value of the nonlocal equivalent strain

$$\tilde{\varepsilon}_{\mathrm{eq}} = \mathbf{N}_\varepsilon \tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^* \tag{B.12}$$

and its corresponding set of weight functions

$$w_\varepsilon = \mathbf{N}_\varepsilon w_\varepsilon^* \tag{B.13}$$

in which $\mathbf{N}_\varepsilon$ is a matrix containing the finite element shape functions that interpolate the nodal values of the nonlocal equivalent strains $\tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^*$ and w_ε^*. By defining $\mathbf{B}_\varepsilon = \nabla \mathbf{N}_\varepsilon$, the weak form (B.7) can be rewritten as

$$\int_\Omega \left((w_\varepsilon^*)^{\mathrm{T}} \mathbf{N}_\varepsilon^{\mathrm{T}} \mathbf{N}_\varepsilon \tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^* + (w_\varepsilon^*)^{\mathrm{T}} \mathbf{B}_\varepsilon^{\mathrm{T}} \mathbf{c} \mathbf{B}_\varepsilon \tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^* \right) \mathrm{d}\Omega = \int_\Omega (w_\varepsilon^*)^{\mathrm{T}} \mathbf{N}_\varepsilon^{\mathrm{T}} \varepsilon_{\mathrm{eq}} \mathrm{d}\Omega \tag{B.14}$$

holding for any arbitrary weight function w_ε^* and therefore yielding

$$\int_\Omega \left(\mathbf{N}_\varepsilon^{\mathrm{T}} \mathbf{N}_\varepsilon \tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^* + \mathbf{B}_\varepsilon^{\mathrm{T}} \mathbf{c} \mathbf{B}_\varepsilon \tilde{\boldsymbol{\varepsilon}}_{\mathrm{eq}}^* \right) \mathrm{d}\Omega = \int_\Omega \mathbf{N}_\varepsilon^{\mathrm{T}} \varepsilon_{\mathrm{eq}} \mathrm{d}\Omega . \tag{B.15}$$

Consistent linearization. To achieve a computational efficient and objective procedure for tracing nonlinear material behavior, a consistent linearization should be considered by deriving proper tangent moduli for the use of these models in incremental-iterative solution algorithms. Moreover, since $\mathbf{c}$ is no longer a constant but depends on the stress field σ that, in turn, depends on the unknown nodal fields $\mathbf{u}^*$ and $\boldsymbol{\varepsilon}_{\mathrm{eq}}^*$, the degree of nonlinearity becomes more complex, motivating even more the use of proper tangential stiffness matrices. From an implementation point of view, one might be tempted to simply replace the scalar c in conventional isotropic gradient-enhanced damage algorithms by the 2×2 gradient matrix $\mathbf{c}$ and keeping it constant during an

50

increment, as done in the M.Sc. thesis of Guido Bongers supervised by the second author and later reproduced in the isogeometric setting by Thai et al. [30, equation (59)]. This approach however may lead to step-size dependent analysis results. As examined in the previous section, there are two coupled field equations that should be simultaneously satisfied:

$$\mathbf{f} = \mathbf{f}_{\text{int}} - \mathbf{f}_{\text{ext}} = \int_{\Omega} \mathbf{B}_u^{\text{T}} \sigma \mathrm{d}\Omega - \int_{\Gamma} \mathbf{N}_u^{\text{T}} \mathbf{t} \mathrm{d}\Gamma = \mathbf{0} \,, \tag{B.16}$$

with $\sigma = \mathbf{C}^{\text{s}} \mathbf{B}_u \mathbf{u}$, and

$$\mathbf{g} = \int_{\Omega} \left(\mathbf{N}_{\varepsilon}^{\text{T}} \mathbf{N}_{\varepsilon} \tilde{\varepsilon}_{\text{eq}} + \mathbf{B}_{\varepsilon}^{\text{T}} \mathbf{c} \mathbf{B}_{\varepsilon} \tilde{\varepsilon}_{\text{eq}} - \mathbf{N}_{\varepsilon}^{\text{T}} \varepsilon_{\text{eq}} \right) \mathrm{d}\Omega = \mathbf{0} \,. \tag{B.17}$$

In both equations, the superscript $*$ that refers to nodal quantities has been dropped for clarity. In a nonlinear material context, displacements $\mathbf{u}$ and nonlocal equivalent strains $\tilde{\varepsilon}_{\text{eq}}$ are linearized at iteration i according to

$$\mathbf{u}^i = \mathbf{u}^{i-1} + \Delta \mathbf{u}^i \tag{B.18}$$

and

$$\tilde{\varepsilon}_{\text{eq}}^i = \tilde{\varepsilon}_{\text{eq}}^{i-1} + \Delta \tilde{\varepsilon}_{\text{eq}}^i \tag{B.19}$$

in which the corrections $\Delta \mathbf{u}^i$ and $\Delta \tilde{\varepsilon}_{\text{eq}}^i$ are the solutions of the coupled system of equations

$$\begin{bmatrix} \mathbf{K}_{uu} & \mathbf{K}_{u\varepsilon} \\ \mathbf{K}_{\varepsilon u} & \mathbf{K}_{\varepsilon\varepsilon} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{u}^i \\ \Delta \tilde{\varepsilon}_{\text{eq}}^i \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\text{ext}} - \mathbf{f}_{\text{int}}^{i-1} \\ -\mathbf{g}^{i-1} \end{bmatrix} \tag{B.20}$$

where the external forces acting on the surface Γ are calculated according to

$$\mathbf{f}_{\text{ext}} = \int_{\Gamma} \mathbf{N}_u^{\text{T}} \mathbf{t} \mathrm{d}\Gamma \,, \tag{B.21}$$

the internal out-of-balance forces equal

$$\mathbf{f}_{\text{int}}^{i-1} = \int_{\Omega} \mathbf{B}_u^{\text{T}} \sigma^{i-1} \mathrm{d}\Omega \,, \tag{B.22}$$

and the out-of-balance right-hand side term related to the equivalent strains, $\mathbf{g}^{i-1}$, is calculated according to (B.17) using the local and nonlocal equivalent strains from the previous iteration $i-1$. The terms of the Jacobian in (B.20) are expressed as

$$\mathbf{K}_{uu} = \frac{\partial \mathbf{f}_{\text{int}}^{i-1}}{\partial \mathbf{u}} \,, \tag{B.23}$$

$$\mathbf{K}_{u\varepsilon} = \frac{\partial \mathbf{f}_{\text{int}}^{i-1}}{\partial \tilde{\varepsilon}_{\text{eq}}} \,, \tag{B.24}$$

$$\mathbf{K}_{\varepsilon u} = \frac{\partial \mathbf{g}^{i-1}}{\partial \mathbf{u}} \,, \tag{B.25}$$

and

$$\mathbf{K}_{\varepsilon\varepsilon} = \frac{\partial \mathbf{g}^{i-1}}{\partial \tilde{\varepsilon}_{\text{eq}}} \,. \tag{B.26}$$

51

Using the chain rule, these terms are expressed as

$$\mathbf{K}_{uu} = \frac{\partial \mathbf{f}_{\text{int}}^{i-1}}{\partial \boldsymbol{\sigma}} \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} \frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{u}} = \int_{\Omega} \mathbf{B}_u^{\text{T}} (1 - \omega) \, \mathbf{C} \mathbf{B}_u \mathrm{d}\Omega \,, \tag{B.27}$$

$$\mathbf{K}_{u\varepsilon} = \frac{\partial \mathbf{f}_{\text{int}}^{i-1}}{\partial \boldsymbol{\sigma}} \frac{\partial \boldsymbol{\sigma}}{\partial \omega} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \frac{\partial \tilde{\varepsilon}_{\text{eq}}}{\partial \tilde{\varepsilon}_{\text{eq}}} = - \int_{\Omega} \mathbf{B}_u^{\text{T}} \mathbf{C} \boldsymbol{\varepsilon} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \mathbf{N}_{\varepsilon} \mathrm{d}\Omega \,, \tag{B.28}$$

$$\mathbf{K}_{\varepsilon u} = \frac{\partial \mathbf{g}^{i-1}}{\partial \varepsilon_{\text{eq}}} \frac{\partial \varepsilon_{\text{eq}}}{\partial \boldsymbol{\varepsilon}} \frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{u}} + \frac{\partial \mathbf{g}^{i-1}}{\partial \mathbf{c}} \frac{\partial \mathbf{c}}{\partial \mathbf{u}} = - \int_{\Omega} \mathbf{N}_{\varepsilon}^{\text{T}} \frac{\partial \varepsilon_{\text{eq}}}{\partial \boldsymbol{\varepsilon}} \mathbf{B}_u \mathrm{d}\Omega + \int_{\Omega} \mathbf{B}_{\varepsilon}^{\text{T}} \frac{\partial \mathbf{c}}{\partial \mathbf{u}} \mathbf{B}_{\varepsilon} \tilde{\varepsilon}_{\text{eq}} \mathrm{d}\Omega \,, \tag{B.29}$$

and

$$\mathbf{K}_{\varepsilon\varepsilon} = \int_{\Omega} \left(\mathbf{N}_{\varepsilon}^{\text{T}} \mathbf{N}_{\varepsilon} + \mathbf{B}_{\varepsilon}^{\text{T}} \mathbf{c} \mathbf{B}_{\varepsilon} \right) \mathrm{d}\Omega + \int_{\Omega} \mathbf{B}_{\varepsilon}^{\text{T}} \frac{\partial \mathbf{c}}{\partial \tilde{\varepsilon}_{\text{eq}}} \mathbf{B}_{\varepsilon} \tilde{\varepsilon}_{\text{eq}} \mathrm{d}\Omega \,, \tag{B.30}$$

where it is noted that the second integral terms in the latter two expressions vanish in case of a constant gradient matrix $\mathbf{c}$. When $\mathbf{c}$ is not constant, the partial derivative $\frac{\partial \mathbf{c}}{\partial \mathbf{u}}$ in (B.29) is calculated as

$$\frac{\partial \mathbf{c}}{\partial \mathbf{u}} = \frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}} \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} \frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{u}} = \frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}} (1 - \omega) \, \mathbf{C} \mathbf{B}_u \tag{B.31}$$

where $\frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}}$ depends on the definition of the gradient matrix $\mathbf{c}$, i.e., (19) and (30). The term $\frac{\partial \mathbf{c}}{\partial \tilde{\varepsilon}_{\text{eq}}}$ in (B.30) is expanded as

$$\frac{\partial \mathbf{c}_{\sigma}}{\partial \tilde{\varepsilon}_{\text{eq}}} = \frac{\partial \mathbf{c}_{\sigma}}{\partial \boldsymbol{\sigma}} \frac{\partial \boldsymbol{\sigma}}{\partial \omega} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \frac{\partial \tilde{\varepsilon}_{\text{eq}}}{\partial \tilde{\varepsilon}_{\text{eq}}} = \frac{\partial \mathbf{c}_{\sigma}}{\partial \boldsymbol{\sigma}} \left(-\mathbf{C}\boldsymbol{\varepsilon} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \mathbf{N}_{\varepsilon} \right) \tag{B.32}$$

and

$$\begin{aligned} \frac{\partial \mathbf{c}_{\tilde{\sigma}}}{\partial \tilde{\varepsilon}_{\text{eq}}} &= cs \frac{\partial \mathbf{c}_{\sigma,\text{norm}}}{\partial \boldsymbol{\sigma}} \frac{\partial \boldsymbol{\sigma}}{\partial \omega} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \frac{\partial \tilde{\varepsilon}_{\text{eq}}}{\partial \tilde{\varepsilon}_{\text{eq}}} + c\mathbf{c}_{\sigma,\text{norm}} \frac{\partial s}{\partial \tilde{\varepsilon}_{\text{eq}}} \\ &= cs \frac{\partial \mathbf{c}_{\sigma,\text{norm}}}{\partial \boldsymbol{\sigma}} \left(-\mathbf{C}\boldsymbol{\varepsilon} \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \mathbf{N}_{\varepsilon} \right) + c\mathbf{c}_{\sigma,\text{norm}} \frac{1}{\kappa_0} \left((1 - \omega) \, \mathbf{N}_{\varepsilon} - \frac{\partial \omega}{\partial \kappa} \frac{\partial \kappa}{\partial \tilde{\varepsilon}_{\text{eq}}} \mathbf{N}_{\varepsilon} \tilde{\varepsilon}_{\text{eq}} \right) \end{aligned} \tag{B.33}$$

for the principal stress-based and nonlocal equivalent stress-based model, respectively. Due to their lengthy expressions, the expanded derivatives $\frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}}$ in (B.31) and (B.32) and $\frac{\partial \mathbf{c}_{\sigma,\text{norm}}}{\partial \boldsymbol{\sigma}}$ in (B.33) are not given here but can be easily computed using a symbolic algebraic software. As discussed in Section 4, $\mathbf{c}$ can be replaced by its filtered value $\mathbf{c}'_{\text{min}}$ to prevent the appearance of positive off-diagonal terms in $\mathbf{K}_{\varepsilon\varepsilon}$. In that case, $\frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}}$ is replaced by

$$\frac{\partial \mathbf{c}'_{\text{min}}}{\partial \mathbf{c}} \frac{\partial \mathbf{c}}{\partial \boldsymbol{\sigma}} \tag{B.34}$$

in the expressions above.

Appendix C. Minimum values of $\mathbf{c}_{\text{min,IP}}$ for T3 elements integrated using a three-point Gauss scheme

To cancel the off-diagonal terms in the integration point stiffness matrices $\mathbf{K}_{\varepsilon\varepsilon,\text{IP}}$, a unique set of terms $c_{xx,\text{min,IP}}$, $c_{yy,\text{min,IP}}$, and $c_{xy,\text{min,IP}}$ can be derived for a three-node triangular element that

52

is numerically integrated using a three-point Gauss quadrature scheme. For the integration point closest to node (x_1, y_1) these terms are

$$c_{xx,\min,\text{IP1}} = \frac{1}{36}\left(8x_1{}^2 + 5x_2{}^2 + 5x_3{}^2 - 8x_1x_2 - 2x_2x_3 - 8x_3x_1\right),\tag{C.1}$$

$$c_{yy,\min,\text{IP1}} = \frac{1}{36}\left(8y_1{}^2 + 5y_2{}^2 + 5y_3{}^2 - 8y_1y_2 - 2y_2y_3 - 8y_3y_1\right),\tag{C.2}$$

and

$$c_{xy,\min,\text{IP1}} = \frac{1}{36}\left(8x_1y_1 + 5x_2y_2 + 5x_3y_3 - 4x_1y_2 - 4x_2y_1 - x_2y_3 - x_3y_2 - 4x_1y_3 - 4x_3y_1\right),\tag{C.3}$$

whereas for the integration point closest to the node (x_2, y_2) the terms of $\mathbf{c}_{\min}$ are given by

$$c_{xx,\min,\text{IP2}} = \frac{1}{36}\left(5x_1{}^2 + 8x_2{}^2 + 5x_3{}^2 - 8x_1x_2 - 8x_2x_3 - 2x_3x_1\right),\tag{C.4}$$

$$c_{yy,\min,\text{IP2}} = \frac{1}{36}\left(5y_1{}^2 + 8y_2{}^2 + 5y_3{}^2 - 8y_1y_2 - 8y_2y_3 - 2y_3y_1\right),\tag{C.5}$$

and

$$c_{xy,\min,\text{IP2}} = \frac{1}{36}\left(5x_1y_1 + 8x_2y_2 + 5x_3y_3 - 4x_1y_2 - 4x_2y_1 - 4x_2y_3 - 4x_3y_2 - x_1y_3 - x_3y_1\right),\tag{C.6}$$

and, finally, for the integration point closest to the node (x_3, y_3) the terms read

$$c_{xx,\min,\text{IP3}} = \frac{1}{36}\left(5x_1{}^2 + 5x_2{}^2 + 8x_3{}^2 - 2x_1x_2 - 8x_2x_3 - 8x_3x_1\right),\tag{C.7}$$

$$c_{yy,\min,\text{IP3}} = \frac{1}{36}\left(5y_1{}^2 + 5y_2{}^2 + 8y_3{}^2 - 2y_1y_2 - 8y_2y_3 - 8y_3y_1\right),\tag{C.8}$$

and

$$c_{xy,\min,\text{IP3}} = \frac{1}{36}\left(5x_1y_1 + 5x_2y_2 + 8x_3y_3 - x_1y_2 - x_2y_1 - 4x_2y_3 - 4x_3y_2 - 4x_1y_3 - 4x_3y_1\right).\tag{C.9}$$

Appendix D. Inequality constants for T3 elements integrated using a three-point Gauss scheme

At integration point level, the sign evaluation of the three unique off-diagonal terms in the symmetric element matrix $\mathbf{K}_{\varepsilon\varepsilon}$, expressed by (46), can be considered as a verification of three inequalities of the form

$$p_{2,1}c_{xx} + q_{2,1}c_{yy} + r_{2,1}c_{xy} + s_{2,1} \leq \rho,\tag{D.1}$$

$$p_{3,1}c_{xx} + q_{3,1}c_{yy} + r_{3,1}c_{xy} + s_{3,1} \leq \rho,\tag{D.2}$$

and

$$p_{3,2}c_{xx} + q_{3,2}c_{yy} + r_{3,2}c_{xy} + s_{3,2} \leq \rho,\tag{D.3}$$

in which the constants $p_{i,j}$, $q_{i,j}$, $r_{i,j}$, and $s_{i,j}$ follow from the (integration point dependent) value of the respective terms of the shape function matrix $\mathbf{N}_{\varepsilon\varepsilon}$ and the matrix containing the partial

derivatives of the shape functions $\mathbf{B}_{\varepsilon\varepsilon}$. For the integration point closest to node (x_1, y_1) these terms are:

$$p_{2,1} = -(y_1 - y_3)(y_2 - y_3)\,, \tag{D.4}$$

$$q_{2,1} = -(x1 - x_3)(x_2 - x_3)\,, \tag{D.5}$$

$$r_{2,1} = x_1 y_2 + x_2 y_1 - x_1 y_3 - x_3 y_1 - x_2 y_3 - x_3 y_2 + 2 x_3 y_3\,, \tag{D.6}$$

$$s_{2,1} = \frac{1}{9}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,, \tag{D.7}$$

$$p_{3,1} = (y_1 - y_2)(y_2 - y_3)\,, \tag{D.8}$$

$$q_{3,1} = (x_1 - x_2)(x_2 - x_3)\,, \tag{D.9}$$

$$r_{3,1} = x_1 y_3 - x_2 y_1 - x_1 y_2 + 2 x_2 y_2 + x_3 y_1 - x_2 y_3 - x_3 y_2\,, \tag{D.10}$$

$$s_{3,1} = \frac{1}{9}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,, \tag{D.11}$$

$$p_{3,2} = -(y_1 - y_2)(y_1 - y_3)\,, \tag{D.12}$$

$$q_{3,2} = -(x_1 - x_2)(x_1 - x_3)\,, \tag{D.13}$$

$$r_{3,2} = 2 x_1 y_1 - x_1 y_2 - x_1 y_3 + x_2 y_3 - x_3 y_1 - x_2 y_1 + x_3 y_2\,, \tag{D.14}$$

and

$$s_{3,2} = \frac{1}{36}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,. \tag{D.15}$$

For the integration point closest to node (x_2, y_2) the corresponding terms equal

$$p_{2,1} = -(y_1 - y_3)(y_2 - y_3)\,, \tag{D.16}$$

$$q_{2,1} = -(x_1 - x_3)(x_2 - x_3)\,, \tag{D.17}$$

$$r_{2,1} = x_1 y_2 + x_2 y_1 - x_1 y_3 - x_3 y_1 - x_2 y_3 - x_3 y_2 + 2 x_3 y_3\,, \tag{D.18}$$

$$s_{2,1} = \frac{1}{9}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,, \tag{D.19}$$

$$p_{3,1} = (y_1 - y_2)(y_2 - y_3)\,, \tag{D.20}$$

$$q_{3,1} = (x_1 - x_2)(x_2 - x_3)\,, \tag{D.21}$$

$$r_{3,1} = x_1 y_3 - x_2 y_1 - x_1 y_2 + 2 x_2 y_2 + x_3 y_1 - x_2 y_3 - x_3 y_2\,, \tag{D.22}$$

$$s_{3,1} = \frac{1}{36}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,, \tag{D.23}$$

$$p_{3,2} = -(y_1 - y_2)(y_1 - y_3)\,, \tag{D.24}$$

$$q_{3,2} = -(x_1 - x_2)(x_1 - x_3)\,, \tag{D.25}$$

$$r_{3,2} = 2 x_1 y_1 - x_1 y_2 - x_1 y_3 + x_2 y_3 - x_3 y_1 - x_2 y_1 + x_3 y_2\,, \tag{D.26}$$

and

$$s_{3,2} = \frac{1}{9}\left(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2\right)^2\,, \tag{D.27}$$

54

whereas for the Gauss point closest to node (x_3, y_3) the constants read

$$p_{2,1} = -(y_1 - y_3)(y_2 - y_3), \tag{D.28}$$

$$q_{2,1} = -(x_1 - x_3)(x_2 - x_3), \tag{D.29}$$

$$r_{2,1} = x_1 y_2 + x_2 y_1 - x_1 y_3 - x_3 y_1 - x_2 y_3 - x_3 y_2 + 2 x_3 y_3, \tag{D.30}$$

$$s_{2,1} = \frac{1}{36}(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2)^2, \tag{D.31}$$

$$p_{3,1} = (y_1 - y_2)(y_2 - y_3), \tag{D.32}$$

$$q_{3,1} = (x_1 - x_2)(x_2 - x_3), \tag{D.33}$$

$$r_{3,1} = x_1 y_3 - x_2 y_1 - x_1 y_2 + 2 x_2 y_2 + x_3 y_1 - x_2 y_3 - x_3 y_2, \tag{D.34}$$

$$s_{3,1} = \frac{1}{9}(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2)^2, \tag{D.35}$$

$$p_{3,2} = -(y_1 - y_2)(y_1 - y_3), \tag{D.36}$$

$$q_{3,2} = -(x_1 - x_2)(x_1 - x_3), \tag{D.37}$$

$$r_{3,2} = 2 x_1 y_1 - x_1 y_2 - x_1 y_3 + x_2 y_3 - x_3 y_1 - x_2 y_1 + x_3 y_2, \tag{D.38}$$

and

$$s_{3,2} = \frac{1}{9}(x_1 y_2 - x_2 y_1 - x_1 y_3 + x_3 y_1 + x_2 y_3 - x_3 y_2)^2. \tag{D.39}$$

References

[1] G. Pijaudier-Cabot and Z. Bažant. Nonlocal damage theory. *Journal of Engineering Mechanics*, 113(10):1512–1533, 1987.

[2] R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans, and J.H.P. De Vree. Gradient enhanced damage for quasi-brittle materials. *International Journal for Numerical Methods in Engineering*, 39(19):3391–3403, 1996.

[3] Z. Bažant, T. Belytschko, and T.P. Chang. Continuum theory for strain-softening. *Journal of Engineering Mechanics*, 110:1666–1692, 1984.

[4] M.G.D. Geers, R. de Borst, W.A.M. Brekelmans, and R.H.J. Peerlings. Strain-based transient-gradient damage model for failure analyses. *Computer Methods in Applied Mechanics and Engineering*, 160(1-2):133–153, 1998.

[5] A. Simone, H. Askes, and L.J. Sluys. Incorrect initiation and propagation of failure in non-local and gradient-enhanced media. *International Journal of Solids and Structures*, 41(2):351–363, 2004.

[6] S. Saroukhani, R. Vafadari, and A. Simone. A simplified implementation of a gradient-enhanced damage model with transient length scale effects. *Computational Mechanics*, 51(6):899–909, 2012.

[7] M.G.D. Geers, R.H.J. Peerlings, W.A.M. Brekelmans, and R. de Borst. Phenomenological nonlocal approaches based on implicit gradient-enhanced damage. *Acta Mechanica*, 144:1–15, 2000.

[8] L.H. Poh and G. Sun. Localizing gradient damage model with decreasing interactions. *International Journal for Numerical Methods in Engineering*, 110(6):503–522, 2016.

[9] G. Pijaudier-Cabot, K. Haidar, and J.-F. Dubé. Non-local damage model with evolving internal length. *International Journal for Numerical and Analytical Methods in Geomechanics*, 28(7–8):633–652, 2004.

[10] G.D. Nguyen. A damage model with evolving nonlocal interactions. *International Journal of Solids and Structures*, 48(10):1544–1559, 2011.

[11] Z. Bažant. Nonlocal damage theory based on micromechanics of crack interaction. *Journal of Engineering Mechanics*, 120:593–617, 1994.

[12] C. Giry, F. Dufour, and J. Mazars. Stress-based nonlocal damage model. *International Journal of Solids and Structures*, 48(25-26):3431–3443, 2011.

[13] L. Wu, F. Sket, J.M. Molina-Aldareguia, A. Makradi, L. Adam, I. Doghri, and L. Noels. A study of composite laminates failure using an anisotropic gradient-enhanced damage mean-field homogenization model. *Composite Structures*, 126:246–264, 2015.

[14] C.V. Verhoosel and R. de Borst. A phase-field model for cohesive fracture. *International Journal for Numerical Methods in Engineering*, 96(1):43–62, 2013.

[15] L. Wu, L. Noels, L. Adam, and I. Doghri. An implicit-gradient-enhanced incremental-secant mean-field homogenization scheme for elasto-plastic composites with damage. *International Journal of Solids and Structures*, 50(24):3843–3860, 2013.

[16] M.A. Gutiérrez. Energy release control for numerical simulations of failure in quasi-brittle solids. *Communications in Numerical Methods in Engineering*, 20(1):19–29, 2003.

[17] S. V. Harren, H. E. Dève, and R. J. Asaro. Shear band formation in plane strain compression. *Acta Metallurgica*, 36(9):2435–2480, 1988.

[18] S. Nemat-Nasser and N. Okada. Radiographic and microscopic observation of shear bands in granular materials. *Géotechnique*, 51(9):753–765, 2001.

[19] A. Rodríguez-Ferran, I. Morata, and A. Huerta. A new damage model based on non-local displacements. *International Journal for Numerical and Analytical Methods in Geomechanics*, 29(5):473–493, 2005.

[20] H. Askes, P. Livieri, L. Susmel, D. Taylor, and R. Tovo. Intrinsic material length, theory of critical distances and gradient mechanics: Analogies and differences in processing linear-elastic crack tip stress fields. *Fatigue & Fracture of Engineering Materials & Structures*, 36(1):39–55, 2012.

[21] A. Simone, G.N. Wells, and L.J. Sluys. From continuous to discontinuous failure in a gradient-enhanced continuum damage model. *Computer Methods in Applied Mechanics and Engineering*, 192(41–42):4581–4607, 2003.

[22] M.A. Celia, E.T. Bouloutas, and R.L. Zarba. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resources Research*, 26(7):1483–1496, 1990.

[23] S.-H. Ju and K.-J. S. Kung. Mass types, element orders and solution schemes for the richards equation. *Computers & Geosciences*, 23(2):175–187, 1997.

[24] H. Ouyang and D. Xiao. Criteria for eliminating oscillation in analysis of heat-conduction equation by finite-element method. *Communications in Numerical Methods in Engineering*, 10:453–460, 1994.

[25] J. Haelewyn. Diagonalization of the thermal mass matrix (CodeAster manual). Technical report, Électricité de France (EDF), 2013.

[26] H. Fujii. Some remarks on finite element analysis of time-dependent field problems. *Theory and Practice in Finite Element Structural Analysis*, pages 91–106, 1973.

[27] E. Rank, C. Katz, and H. Werner. On the importance of the discrete maximum principle in transient analysis using finite element methods. *International Journal for Numerical Methods in Engineering*, 19:1771–1782, 1983.

[28] O. Axelsson and L. Kolotilina. Diagonally compensated reduction and related preconditioning methods. *Numerical Linear Algebra with Applications*, 1(2):155–177, 1994.

[29] J.G.M. van Mier. *Fracture Processes of Concrete*. CRC Press, 1996.

[30] T.Q. Thai, T. Rabczuk, Y. Bazilevs, and G. Meschke. A higher-order stress-based gradient-enhanced damage model based on isogeometric analysis. *Computer Methods in Applied Mechanics and Engineering*, 304:584–604, 2016.