INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2006; **67**:1531–1564
Published online 30 March 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme.1673

Capturing strain localization in dense sands with random density

José E. Andrade and Ronaldo I. Borja*, †

Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, U.S.A.

SUMMARY

This paper presents a three-invariant constitutive framework suitable for the numerical analyses of localization instabilities in granular materials exhibiting unstructured random density. A recently proposed elastoplastic model for sands based on critical state plasticity is enhanced with the third stress invariant to capture the difference in the compressive and extensional yield strengths commonly observed in geomaterials undergoing plastic deformation. The new three-invariant constitutive model, similar to its two-invariant predecessor, is capable of accounting for meso-scale inhomogeneities as well as material and geometric nonlinearities. Details regarding the numerical implementation of the model into a fully nonlinear finite element framework are presented and a closed-form expression for the consistent tangent operator, whose spectral form is used in the strain localization analyses, is derived. An algorithm based on the spectral form of the so-called acoustic tensor is proposed to search for the necessary conditions for deformation bands to develop. The aforementioned framework is utilized in a series of boundary-value problems on dense sand specimens whose density fields are modelled as exponentially distributed unstructured random fields to account for the effect of inhomogeneities at the meso-scale and the intrinsic uncertainty associated with them. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: strain localization; random density; granular materials

1. INTRODUCTION

Shear bands in granular materials, characterized by thin zones of intense shearing, have been a subject of considerable research interest since they impact many problems in the fields of geoengineering and geoscience. In geoengineering, shear bands control the deformation behaviour and stability of earth structures such as dams, tunnels, excavations and foundations. In geoscience, they are commonly associated with grain fracturing and grain size reduction (termed 'cataclasis' in the literature [1]) that are known to alter the strength and transport properties of

*Correspondence to: Ronaldo I. Borja, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, U.S.A.
†E-mail: borja@stanford.edu

Contract/grant sponsor: U.S. National Science Foundation; contract/grant numbers: CMS-0201317, CMS-0324674

Received 4 August 2005
Revised 19 January 2006
Accepted 20 January 2006

geological systems. Various experimental and theoretical studies have thus attempted to analyse and describe the occurrence and patterns of shear bands in granular materials such as soils and rocks.

In sand bodies, shear band formations are now investigated in the laboratory using advanced testing techniques such as gamma-rays, stereophotogrammetry, and X-ray computed tomography (CT), see Reference [2]. Along with the more traditional laboratory testing procedures such as triaxial and simple shear testing, high-end testing provides an opportunity to view the specimen as an imperfect structure. For example, it is now possible to measure quantitatively the density and displacement fields within the specimen with a high degree of precision. Instead of viewing the entire specimen as a homogeneous element, which has been traditionally done in the past, it is now possible to digitally measure and quantify the properties and mechanical responses of an imperfect specimen on a finer scale.

Ironically, whereas development of very sophisticated constitutive models for element testing has previously outpaced the development of laboratory testing capabilities supporting the demands of these complex models, the reverse is now true with the advent of high-end precision laboratory testing. There is now a shortage of mathematical models that can adequately capture the quantified heterogeneities in the properties and mechanical responses of imperfect soil samples. This is because many of these sophisticated models are simply too complex to be implemented efficiently into multi-purpose finite element codes, and so only a handful of these models have been successful in finding their way into general-purpose computer codes. Since heterogeneous samples require that their responses be analysed as a boundary-value problem and not as an elemental point, many of these complex models have to be reformulated first before they can be used for general-purpose boundary-value problem simulations.

A promising approach to simulating material heterogeneity in specimens of discrete granular materials such as sand bodies is through meso-scale modelling. By 'meso-scale' we mean a scale larger than a particle size but smaller than the specimen dimensions. We do not wish to go to particle level at this point since there is currently a lack of testing capabilities to capture all the particle information such as shapes, dimensions, and motion of the individual particles throughout testing. In contrast, some of the high-end testing techniques mentioned above, specifically the CT digital imaging technique, can very well be combined with existing computer tools that permit an accurate quantification of material heterogeneity in soil samples. Examples of such techniques include digital image processing (DIP) for quantifying the spatial density variation and digital image correlation (DIC) for tracking the motion of a group of particles contained in a 'meso-element' having the size of a pixel [3, 4].

From a modelling standpoint, a key variable that quantitatively describes material heterogeneity on a meso-scale level is density. Whereas it may be argued that other constitutive variables such as strength and stiffness could also vary spatially within a soil sample, and therefore they should also be considered as meso-variables, it is generally difficult if not impossible to measure their spatial variation in a heterogeneous soil sample. In contrast, DIP can be used to analyse pixel patterns provided by CT images to correlate the grey level with density pattern in the individual pixels. Furthermore, strength and stiffness are known to correlate strongly with density in soils, so it seems plausible to claim that density is indeed a key measure of material heterogeneity.

Having chosen density as a key measure of material heterogeneity, we now consider a constitutive model based on critical state plasticity theory that permits the specification of spatial density variation within a soil sample. One such model, derived from the 'Nor-Sand'

of Jefferies [5] and presented by Borja and Andrade [6], utilizes a variable called 'state parameter' ψ to describe the density of a meso-element relative to that at the critical state line for the same effective mean normal stress [5, 7, 8]. Where the point lies below the critical state line, ψ is negative (denser than critical), and where it lies above, ψ is positive (looser than critical). Traditional Cam–Clay-type models [9–14], do not have the capability to independently prescribe the density (or void ratio) of a meso-element, since, by prescribing the degree of overconsolidation [9] and the current state of stress, one is bound to calculate a unique void ratio. By introducing the state parameter in the constitutive description, the density (or other equivalent measures thereof, such as the void ratio and specific volume) is effectively 'detached' from the critical state line.

This paper moves one step further in refinement to the two-invariant model presented by Borja and Andrade [6] in that we now also quantify the influence of the third stress invariant in the constitutive description of meso-element behaviour. Granular materials such as soils are known to exhibit yield stresses in compression that are higher than those in extension, so it only makes sense to include the third stress invariant in the constitutive formulation. Furthermore, it has been shown that the third stress invariant does enhance the onset of strain localization [15, 16], thus further motivating the present effort. For the present paper, we introduce the third stress invariant through Lode's angle [17] and employ a return mapping algorithm in principal elastic logarithmic stretches (within the framework of multiplicative plasticity) to carry out the numerical integration in discrete load steps.

Apart from the inclusion of the third stress invariant, the present paper also models the effect of unstructured random density fields in the mathematical characterization of nearly homogeneous soil samples at meso-scale. Borja and Andrade [6] have shown that a well defined density structure, generated deterministically, could greatly impact the position of the resulting shear band. However, in reality most laboratory testing procedures attempt to generate as nearly uniform soil samples as possible. Thus the numerical simulations presented in this paper focus on nearly homogeneous density fields, generated randomly and without any preferred structure, to better capture reality. We show with extensive numerical simulations that even with such minute perturbations in the density field a shear band would still form even in cases where it would not form in perfectly homogeneous specimens. Further, we show that different shear bands could form in each randomly generated specimen, even with a very tight density distribution, affirming the common notion that each physical specimen is unique.

Evaluation of the consistent tangent operator (CTO) is critical for the success of a nonlinear iterative algorithm based on Newton's method, and in this paper we derive such tangent operator in spectral form based on Lie derivatives. Quite recently, Borja [16] has shown that such operator can also be used in lieu of the constitutive tangent operator for strain localization analysis, so the CTO serves a dual role in the present numerical simulations. All simulations presented in this paper have been conducted in the 3D finite deformation regime, including the calculation of the minimum values of the localization function for shear band analysis [18]. For the latter analysis we also present an algorithm for a 3D search in spectral directions.

As for notations and symbols used in this paper, bold-faced letters denote tensors and vectors; the symbol '$\cdot$' denotes an inner product of two vectors (e.g. $\mathbf{a}\cdot\mathbf{b}=a_i b_i$), or a single contraction of adjacent indices of two tensors (e.g. $\mathbf{c}\cdot\mathbf{d}=c_{ij}d_{jk}$); the symbol ':' denotes an inner product of two second-order tensors (e.g. $\mathbf{c}:\mathbf{d}=c_{ij}d_{ij}$), or a double contraction of adjacent indices of tensors of rank two and higher (e.g. $\mathbf{C}:\boldsymbol{\varepsilon}^{\mathrm{e}}=C_{ijkl}\varepsilon^{\mathrm{e}}_{kl}$); the symbol '$\otimes$' denotes a juxtaposition,

e.g. $(\mathbf{a}\otimes\mathbf{b})_{ij} = a_i b_j$. Finally, for any symmetric second-order tensors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, $(\boldsymbol{\alpha}\otimes\boldsymbol{\beta})_{ijkl} = \alpha_{ij}\beta_{kl}$, $(\boldsymbol{\alpha}\oplus\boldsymbol{\beta})_{ijkl} = \beta_{ik}\alpha_{jl}$, and $(\boldsymbol{\alpha}\ominus\boldsymbol{\beta})_{ijkl} = \alpha_{il}\beta_{jk}$.

2. CONSTITUTIVE ASSUMPTIONS

In this section, we present a three-invariant enhancement of the critical state plasticity model by Borja and Andrade [6]. As mentioned in Section 1, geomaterials exhibit a higher yield strength in compression than in extension, so the addition of the third stress invariant is critical for modelling the behaviour of geomaterials more realistically. Further, the hyperelastoplastic model hinges on the well established assumption that the deformation gradient tensor can be decomposed multiplicatively into elastic and plastic parts [19],

$$\mathbf{F} = \mathbf{F}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{p}} \tag{1}$$

where $\mathbf{F}^{\mathrm{e}}$ and $\mathbf{F}^{\mathrm{p}}$ are defined as the elastic and plastic deformation gradient, respectively.

2.1. The hyperelastic model

Consider an isotropic hyperelastic response entailing a strain energy that is a function of either the elastic right Cauchy–Green deformation tensor $\mathbf{C}^{\mathrm{e}}$ or the left Cauchy–Green tensor $\mathbf{b}^{\mathrm{e}}$ i.e. $\Psi = \Psi(\mathbf{C}^{\mathrm{e}}) = \Psi(\mathbf{b}^{\mathrm{e}})$ [20, 21], where

$$\mathbf{C}^{\mathrm{e}} := \mathbf{F}^{\mathrm{et}} \cdot \mathbf{F}^{\mathrm{e}} \quad \text{and} \quad \mathbf{b}^{\mathrm{e}} := \mathbf{F}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{et}} \tag{2}$$

The principal elastic stretches emanate from $\mathbf{F}^{\mathrm{e}} \cdot \mathbf{N}^a = \lambda_a^{\mathrm{e}} \mathbf{n}^a$ (no sum), where λ_a^{e} for $a = 1, 2, 3$ are the principal elastic stretches in the corresponding principal directions $\mathbf{N}^a$ and $\mathbf{n}^a$ in the intermediate and current configuration, respectively. We recall the well known spectral decomposition,

$$\mathbf{b}^{\mathrm{e}} = \sum_{a=1}^{3} \lambda_a^{\mathrm{e}\,2} \mathbf{n}^a \otimes \mathbf{n}^a \tag{3}$$

The elastic region is assumed to be governed by the isotropic strain energy function proposed in Reference [22] and utilized in modelling of granular bodies in References [12, 23]

$$\Psi(\varepsilon_{\mathrm{v}}^{\mathrm{e}}, \varepsilon_{\mathrm{s}}^{\mathrm{e}}) = \tilde{\Psi}(\varepsilon_{\mathrm{v}}^{\mathrm{e}}) + \tfrac{3}{2}\mu^{\mathrm{e}} \varepsilon_{\mathrm{s}}^{\mathrm{e}2} \tag{4}$$

where

$$\tilde{\Psi}(\varepsilon_{\mathrm{v}}^{\mathrm{e}}) = -p_0 \hat{\kappa} \exp \omega, \quad \omega = -\frac{\varepsilon_{\mathrm{v}}^{\mathrm{e}} - \varepsilon_{\mathrm{v}0}^{\mathrm{e}}}{\hat{\kappa}}, \quad \mu^{\mathrm{e}} = \mu_0 + \frac{\alpha_0}{\hat{\kappa}} \tilde{\Psi}(\varepsilon_{\mathrm{v}}^{\mathrm{e}}) \tag{5}$$

The independent variables are the volumetric and deviatoric invariants of the elastic logarithmic stretch tensor, respectively,

$$\varepsilon_{\mathrm{v}}^{\mathrm{e}} = \varepsilon_1^{\mathrm{e}} + \varepsilon_2^{\mathrm{e}} + \varepsilon_3^{\mathrm{e}} \quad \text{and} \quad \varepsilon_{\mathrm{s}}^{\mathrm{e}} = \tfrac{1}{3}\sqrt{2[(\varepsilon_1^{\mathrm{e}} - \varepsilon_2^{\mathrm{e}})^2 + (\varepsilon_2^{\mathrm{e}} - \varepsilon_3^{\mathrm{e}})^2 + (\varepsilon_3^{\mathrm{e}} - \varepsilon_1^{\mathrm{e}})^2]} \tag{6}$$

where $\varepsilon_a^{\mathrm{e}} \equiv \ln \lambda_a^{\mathrm{e}}$. Hence, the strain energy function is an invariant function of the elastic deformations. The Kirchhoff stress tensor $\boldsymbol{\tau}$ is coaxial with the deformation tensor $\mathbf{b}^{\mathrm{e}}$ and

defined such that

$$\boldsymbol{\tau} = 2\frac{\partial \Psi}{\partial \mathbf{b}^{\mathrm{e}}} \cdot \mathbf{b}^{\mathrm{e}} \tag{7}$$

The above hyperelastic model produces pressure-dependent elastic bulk and shear moduli, a feature commonly observed in the laboratory. The elastic constants necessary for a full description of the elasticity are the reference strain $\varepsilon^{\mathrm{e}}_{\mathrm{v}0}$ and the reference pressure p_0 of the elastic compression curve, as well as the compressibility index $\hat{\kappa}$. The model produces coupled volumetric and deviatoric responses in the case $\alpha_0 \neq 0$ for which μ^{e} is a nonlinear function of the volumetric deformations. Otherwise, for $\alpha_0 = 0$ the responses are decoupled and the shear modulus $\mu^{\mathrm{e}} \equiv \mu_0$ is constant.

2.2. Yield surface, plastic potential, their derivatives and the flow rule

We define the three invariants of the Kirchhoff stress tensor as

$$p = \frac{1}{3}\mathrm{tr}\,\boldsymbol{\tau}, \quad q = \sqrt{\frac{3}{2}}\|\boldsymbol{\xi}\|, \quad \frac{1}{\sqrt{6}}\cos 3\theta = \frac{\mathrm{tr}\,\boldsymbol{\xi}^3}{\chi^3} \equiv y \tag{8}$$

where $\boldsymbol{\xi} = \boldsymbol{\tau} - p\mathbf{1}$ is the deviatoric component of the stress tensor $\boldsymbol{\tau}$, and $\chi = \sqrt{\mathrm{tr}\,\boldsymbol{\xi}^2}$. The quantity p is called the mean normal stress and is assumed negative throughout. Further, θ is the Lode's angle whose values range from $0 \leqslant \theta \leqslant \pi/3$; it defines an angle on a deviatoric plane emanating from a tension corner.

From these three stress invariants, we construct a yield surface of the form

$$F(\boldsymbol{\tau}, \pi_{\mathrm{i}}) = F(p, q, \theta, \pi_{\mathrm{i}}) = \zeta(\theta)q + p\eta(p, \pi_{\mathrm{i}}) \tag{9}$$

where

$$\eta = \begin{cases} M[1 + \ln(\pi_{\mathrm{i}}/p)] & \text{if } N = 0 \\ M/N[1 - (1 - N)(p/\pi_{\mathrm{i}})^{N/(1-N)}] & \text{if } N > 0 \end{cases} \tag{10}$$

The image stress $\pi_{\mathrm{i}} < 0$ controls the size of the yield surface; it is defined such that the stress ratio $\eta = -\zeta q/p = M$ when $p = \pi_{\mathrm{i}}$. Note the subscript 'i' stands for image and should not be confused with index notation. The parameter $N \geqslant 0$ determines the curvature of the yield surface on a meridian plane and it typically has a value less than or equal to 0.4 for sands [5]. Lode's angle θ plays the role of the third stress invariant modifying the shape of the yield surface on a deviatoric plane through the function $\zeta = \zeta(\theta)$. Many forms for the function $\zeta(\theta)$ have been proposed in the literature (see Reference [24] for a historical survey). Here we adopt the form proposed by Gudehus [25] and Argyris *et al.* [26], namely,

$$\zeta(\theta, \rho) = \frac{(1 + \rho) + (1 - \rho)\cos 3\theta}{2\rho} \tag{11}$$

where ρ is a constant parameter called ellipticity. The above function is only convex for $\frac{7}{9} \leqslant \rho \leqslant 1$ [27] and satisfies the boundary conditions: (a) $\zeta = 1/\rho$ when $\theta = 0$ i.e. tension corner; and (b) $\zeta = 1$ when $\theta = \pi/3$ i.e. compression corner. A typical three-invariant yield surface of the form presented above for $\rho = 0.78$ is shown in Figure 1. It can be seen from this figure that the yield surface exhibits a greater yield strength under compression than under extension.

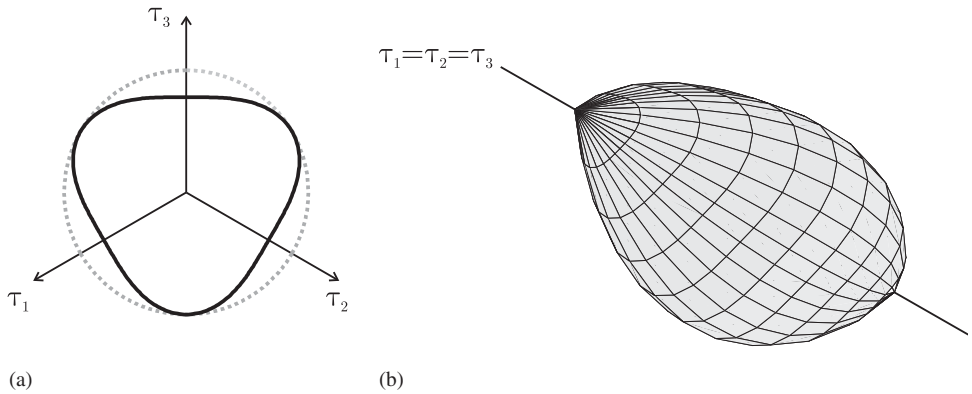

Figure 1. Three-invariant yield surface in Kirchhoff stress space for $\rho = 0.78$: (a) cross-section on deviatoric plane, dashed line represents two-invariant counterpart for comparison; and (b) three-dimensional view.

Additionally, to underscore the fact that the formulation presented herein is general and independent of the choice of the shape function $\zeta(\theta, \rho)$, in some of our numerical examples, we will also utilize the form proposed by Willam and Warnke [28] i.e.

$$\zeta(\theta, \rho) = \frac{4(1-\rho^2)\cos^2\theta + (2\rho-1)^2}{2(1-\rho^2)\cos\theta + (2\rho-1)[4(1-\rho^2)\cos^2\theta + 5\rho^2 - 4\rho]^{1/2}} \tag{12}$$

This function is smooth and convex in the range $\frac{1}{2} \leqslant \rho \leqslant 1$, and has the same boundary conditions than the shape function in (11).

Similar to the yield surface, we can postulate a plastic potential function of the form

$$Q(\boldsymbol{\tau}, \overline{\pi}_{\mathrm{i}}) = Q(p, q, \theta, \overline{\pi}_{\mathrm{i}}) = \overline{\zeta}(\theta) q + p\overline{\eta}(p, \overline{\pi}_{\mathrm{i}}) \tag{13}$$

with

$$\overline{\eta} = \begin{cases} M[1 + \ln(\overline{\pi}_{\mathrm{i}}/p)] & \text{if } \overline{N} = 0 \\ M/\overline{N}[1 - (1-\overline{N})(p/\overline{\pi}_{\mathrm{i}})^{\overline{N}/(1-\overline{N})}] & \text{if } \overline{N} > 0 \end{cases} \tag{14}$$

When $\overline{\pi}_{\mathrm{i}} = \pi_{\mathrm{i}}$, $\overline{N} = N$ and $\overline{\zeta} = \zeta$, plastic flow is associative; otherwise, it is nonassociative in both the volumetric and deviatoric sense. The material parameter $\overline{N}$ controls the amount of *volumetric* nonassociativity, whereas the shape function $\overline{\zeta} = \zeta(\theta, \overline{\rho})$ plays a similar role than that used in the yield surface, but the ellipticity $\overline{\rho}$ can be different from ρ, hence introducing *deviatoric* nonassociativity. We note in passing that the plastic potential in (13) is essentially that proposed by Jefferies [5] for $\overline{\zeta} = 1$.

Recall the multiplicative decomposition of the deformation gradient tensor in (1), which leads to an additive decomposition of the velocity gradient $\mathbf{l}$,

$$\mathbf{l} = \mathbf{l}^{\mathrm{e}} + \mathbf{l}^{\mathrm{p}} \;\Rightarrow\; \mathbf{d} = \mathbf{d}^{\mathrm{e}} + \mathbf{d}^{\mathrm{p}} \tag{15}$$

where $\mathbf{d} \equiv \operatorname{sym} \mathbf{l}$, $\mathbf{d}^{\mathrm{e}} \equiv \operatorname{sym} \mathbf{l}^{\mathrm{e}}$, and $\mathbf{d}^{\mathrm{p}} \equiv \operatorname{sym} \mathbf{l}^{\mathrm{p}}$. Neglecting the plastic spin $\boldsymbol{\omega}^{\mathrm{p}}$ (see Reference [29] for significance and consequences), we write the flow rule as

$$\mathbf{d}^{\mathrm{p}} = \dot{\lambda} \mathbf{q} \tag{16}$$

where $\dot{\lambda}$ is the so-called consistency parameter, and $\mathbf{q} := \partial Q / \partial \boldsymbol{\tau}$. Adopting the spectral approach [15], we recast the above gradient as

$$\mathbf{q} = \frac{\partial Q}{\partial \boldsymbol{\tau}} = \sum_{a=1}^{3} \frac{\partial Q}{\partial \tau_a} \mathbf{m}^a = \sum_{a=1}^{3} q_a \mathbf{m}^a \tag{17}$$

where $\mathbf{m}^a := \mathbf{n}^a \otimes \mathbf{n}^a$, and the vectors $\mathbf{n}^a$ for $a = 1, 2, 3$, are the principal directions of the stress tensor $\boldsymbol{\tau}$, i.e. $\boldsymbol{\tau} \cdot \mathbf{n}^a = \tau_a \mathbf{n}^a$, with principal stresses τ_a for $a = 1, 2, 3$. Note that the stress tensor $\boldsymbol{\tau}$ and the elastic left Cauchy–Green deformation tensor $\mathbf{b}^{\mathrm{e}}$ are coaxial due to the isotropy assumption made in Section 2.1. The spectral decomposition of the flow rule is consequently a by-product of the coaxiality of the tensors $\boldsymbol{\tau}$ and $\mathbf{b}^{\mathrm{e}}$ as well as the fact that Q is an isotropic function of the stress tensor and hence its gradient and the stress tensor are also coaxial. It is then useful to expand q_a such that

$$q_a = \frac{\partial Q}{\partial p} \frac{\partial p}{\partial \tau_a} + \frac{\partial Q}{\partial q} \frac{\partial q}{\partial \tau_a} + \frac{\partial Q}{\partial \theta} \frac{\partial \theta}{\partial \tau_a} \tag{18}$$

where we can easily show that

$$\frac{\partial p}{\partial \tau_a} = \frac{1}{3} \delta_a, \quad \frac{\partial q}{\partial \tau_a} = \sqrt{\frac{3}{2}} \frac{\xi_a}{\chi} \equiv \sqrt{\frac{3}{2}} \hat{n}_a, \quad \frac{\partial \theta}{\partial \tau_a} \equiv \theta_a = -\left(\frac{2}{\sqrt{6}} \csc 3\theta \right) y_a \tag{19}$$

where $\delta_a = 1$, and ξ_a for $a = 1, 2, 3$ are the eigenvalues of the deviatoric stress tensor $\boldsymbol{\xi}$. Additionally,

$$y_a = \frac{\partial y}{\partial \tau_a} = 3 \frac{\xi_a^2}{\chi^3} - 3 \frac{\operatorname{tr}(\boldsymbol{\xi}^3) \xi_a}{\chi^5} - \frac{\delta_a}{\chi} \tag{20}$$

which we use to write q_a in more compact form

$$q_a = \tfrac{1}{3} \partial_p Q \delta_a + \sqrt{\tfrac{3}{2}} \partial_q Q \hat{n}_a + \partial_\theta Q \theta_a \tag{21}$$

From Equations (9) and (13), we obtain

$$\frac{\partial Q}{\partial p} = \beta \frac{\partial F}{\partial p}, \quad \frac{\partial Q}{\partial q} = \bar{\zeta}, \quad \frac{\partial Q}{\partial \theta} = \bar{\zeta}' q \tag{22}$$

where we have used the volumetric nonassociativity parameter $\beta := (1 - N)/(1 - \overline{N})$. Thus,

$$q_a = \tfrac{1}{3} \beta \partial_p F \delta_a + \sqrt{\tfrac{3}{2}} \partial_q Q \hat{n}_a + \partial_\theta Q \theta_a \tag{23}$$

from where we clearly see the volumetric and deviatoric nonassociativity in the plastic flow. Using the above expression for the yield surface, we calculate the derivatives

$$\frac{\partial F}{\partial p} = \frac{\eta - M}{1 - N}, \quad \frac{\partial F}{\partial q} = \zeta, \quad \frac{\partial F}{\partial \theta} = \zeta' q \tag{24}$$

It is now convenient to evaluate the second derivative of the plastic potential with respect to the stress tensor, as it will be used in the numerical implementation of the model. Here, we exploit the uncoupling between the volumetric and deviatoric terms in F and write

$$q_{ab} = \tfrac{1}{9}\beta\partial^2_{pp}F\delta_a\delta_b + \sqrt{\tfrac{3}{2}}\partial_q Q\hat{n}_{ab} + \partial_\theta Q\theta_{ab} + \partial^2_{\theta\theta}Q\theta_a\theta_b + \sqrt{\tfrac{3}{2}}\partial^2_{q\theta}Q(\hat{n}_a\theta_b + \theta_a\hat{n}_b) \quad (25)$$

where from Equations (24) and (22), we obtain

$$\partial^2_{pp}F = -\frac{M}{1-N}\frac{1}{p}\left(\frac{p}{p_\mathrm{i}}\right)^{N/(1-N)}, \quad \partial^2_{\theta\theta}Q = \overline{\zeta}''q, \quad \partial^2_{q\theta}Q = \overline{\zeta}' \quad (26)$$

Furthermore, from Equations $(19)_{2,3}$ we obtain, respectively,

$$\hat{n}_{ab} = \frac{1}{\chi}\left(\delta_{ab} - \frac{1}{3}\delta_a\delta_b - \hat{n}_a\hat{n}_b\right) \quad \text{and} \quad \theta_{ab} = -\left(\frac{2}{\sqrt{6}}\csc 3\theta\right)y_{ab} - (3\cot 3\theta)\theta_a\theta_b \quad (27)$$

with

$$y_{ab} = 6\frac{\xi_a\delta_{ab}}{\chi^3} - 3\frac{\operatorname{tr}\boldsymbol{\xi}^3}{\chi^5}\left(\delta_{ab} - \frac{1}{3}\delta_a\delta_b - 5\frac{\xi_a\xi_b}{\chi^2}\right) + \frac{1}{\chi^3}(\delta_a\xi_b + \xi_a\delta_b) - \frac{9}{\chi^5}(\xi_a\xi_b^2 + \xi_a^2\xi_b) \quad \text{(no sum)} \quad (28)$$

2.2.1. Reduced dissipation inequality: *the second law of thermodynamics*. For perfect plasticity, the reduced dissipation inequality requires the stresses to perform nonnegative plastic incremental work [30], i.e.

$$\mathscr{D}^\mathrm{p} \equiv \boldsymbol{\tau} : \mathbf{d}^\mathrm{p} = \dot{\lambda}\boldsymbol{\tau} : \mathbf{q} = \dot{\lambda}\sum_{a=1}^{3}\tau_a q_a \geqslant 0 \quad (29)$$

Realizing that $\sum_{a=1}^{3}\tau_a\theta_a = 0$, we can write

$$\mathscr{D}^\mathrm{p} = \dot{\lambda}(p\partial_p Q + q\partial_q Q) = -\dot{\lambda}p\left(\frac{M-\overline{\eta}}{1-\overline{N}} + \overline{\eta}\right) \geqslant 0 \quad (30)$$

where we have exploited the fact that at yield $F = Q = 0$. Since $-\dot{\lambda}p \geqslant 0$ and $1 - \overline{N} > 0$, we require $-\overline{N}\overline{\eta} + M \geqslant 0$. Now, for a stress point on the yield surface, the relationship $\overline{\eta}/\overline{\zeta} = \eta/\zeta$ holds, hence,

$$\mathscr{D}^\mathrm{p} \geqslant 0 \;\Rightarrow\; -M\frac{\overline{N}}{N}\frac{\overline{\zeta}}{\zeta}\left[1 - (1-N)\left(\frac{p}{\pi_\mathrm{i}}\right)^{N/(1-N)}\right] + M \geqslant 0 \quad (31)$$

It can be seen that the function inside the square brackets is a monotonically decreasing function and always less than unity for $p \in [\pi_\mathrm{c}, 0]$, where $\pi_\mathrm{c} = \pi_\mathrm{i}/(1-N)^{(1-N)/N}$ is the preconsolidation pressure. Therefore, since $M > 0$ as it is a physical parameter, we get the condition $(\overline{N}\,\overline{\zeta})/(N\zeta) \leqslant 1$, which is satisfied at all times if

$$\overline{N} \leqslant N \quad \text{and} \quad \overline{\zeta} \leqslant \zeta \quad (32)$$

The first condition was already derived by the authors in Reference [6] in the context of two-invariant plasticity, whereas the second condition is a by-product of the third invariant. Both restrictions derived above can be linked to the angle of friction and angle of dilation as $\overline{N} \leqslant N$ implies lower dilatancy than that achieved from a volumetric associative flow rule. On the other hand, it is easy to show that the second condition requires $\rho \leqslant \overline{\rho}$, and since

$$\rho \equiv \frac{3 - \sin \phi_{\mathrm{c}}}{3 + \sin \phi_{\mathrm{c}}} \quad \text{and} \quad \overline{\rho} \equiv \frac{3 - \sin \psi_{c}}{3 + \sin \psi_{\mathrm{c}}} \tag{33}$$

where ϕ_{c} is the angle of internal friction at critical and ψ_{c} is the dilatancy angle at critical, it follows that

$$\psi_{\mathrm{c}} \leqslant \phi_{\mathrm{c}} \tag{34}$$

which is very much consistent with thermodynamic conditions found for Mohr–Coulomb or Drucker–Prager materials and experimental observations in geomaterials, see References [15, 31] for further elaboration.

Remark 1
Experimental evidence in granular materials seems to suggest deviatoric nonassociativity in these materials is not pronounced. Lade and Duncan [32] have shown that, for dense sands, the plastic strain rates are perpendicular to the yield surface on a deviatoric plane. Nevertheless, the formulation presented above is general and allows for deviatoric and volumetric nonassociativity.

2.3. Maximum plastic dilatancy, hardening law and the consistency condition

We recall the definitions for the volumetric and deviatoric plastic strain rate invariants, respectively,

$$\dot{\varepsilon}^{\mathrm{p}}_{\mathrm{v}} = \operatorname{tr} \mathbf{d}^{\mathrm{p}} \quad \text{and} \quad \dot{\varepsilon}^{\mathrm{p}}_{\mathrm{s}} = \sqrt{\tfrac{2}{3}} \| \mathbf{d}^{\mathrm{p}} - \tfrac{1}{3} \dot{\varepsilon}^{\mathrm{p}}_{\mathrm{v}} \mathbf{1} \| \tag{35}$$

Using the definitions above and Equations (16)–(17), and (21) we obtain

$$\dot{\varepsilon}^{\mathrm{p}}_{\mathrm{v}} = \dot{\lambda} \sum_{a=1}^{3} q_a \quad \text{and} \quad \dot{\varepsilon}^{\mathrm{p}}_{\mathrm{s}} = \dot{\lambda} \sqrt{\frac{2}{3} \sum_{a=1}^{3} \tilde{q}_a^2} \tag{36}$$

where we have $\sum_{a=1}^{3} q_a = \partial_p Q$, and $\tilde{q}_a = q_a - \frac{1}{3} \partial_p Q \delta_a$. Hence,

$$\dot{\varepsilon}^{\mathrm{p}}_{\mathrm{v}} = \dot{\lambda} \beta \partial_p F \quad \text{and} \quad \dot{\varepsilon}^{\mathrm{p}}_{\mathrm{s}} = \dot{\lambda} \sqrt{\tfrac{2}{3}} \Omega(q, \theta) \tag{37}$$

with

$$\Omega(\boldsymbol{\tau}) = \Omega(q, \theta) = \sqrt{\tfrac{3}{2} (\partial_q Q)^2 + (\partial_\theta Q)^2 [\theta_1^2 + \theta_2^2 + \theta_3^2]} \tag{38}$$

Now, we recall the definition of plastic dilatancy

$$D := \frac{\dot{\varepsilon}^{\mathrm{p}}_{\mathrm{v}}}{\dot{\varepsilon}^{\mathrm{p}}_{\mathrm{s}}} = \sqrt{\frac{3}{2}} \beta \frac{\partial_p F}{\Omega} \tag{39}$$

which we can use to obtain the maximum dilatancy and thereby maximum stress ratio η^* such that

$$D^* = \alpha\psi_{\mathrm{i}} = \sqrt{\frac{3}{2}}\beta\frac{\eta^* - M}{(1-N)\Omega} \tag{40}$$

After rearranging, we get $\eta^* = \sqrt{\frac{2}{3}}\overline{\alpha}\psi_{\mathrm{i}}(1-N)\Omega + M$, with $\alpha = \beta\overline{\alpha}$ and $\alpha \approx -3.5$ for sands. Recall the maximum plastic dilatancy is used to limit the amount of dilation in a sand specimen on the 'wet' side of the critical state line. As in the original model proposed by Been and Jefferies [7] and Jefferies [5], the maximum dilatancy is obtained through an empirical relation with the state parameter ψ_{i}, which is a distance between the specific volume of the sample and the specific volume at critical at the image pressure i.e.

$$\psi_{\mathrm{i}} = v - v_{\mathrm{c0}} + \hat{\lambda}\ln(-\pi_{\mathrm{i}}) \tag{41}$$

where v is the specific volume, v_{c0} is the reference specific volume at unit pressure, and $\hat{\lambda}$ is the plastic compressibility index. All of these parameters emanate from the so-called critical state theory which postulates the existence of the critical state line. Inserting the above result into Equation (10) and solving for the corresponding limiting image pressure we get, cf. Reference [6],

$$\frac{\pi_{\mathrm{i}}^*}{p} = \begin{cases} \exp(\sqrt{\frac{2}{3}}\overline{\alpha}/M\psi_{\mathrm{i}}\Omega) & \text{if } \overline{N} = N = 0 \\ (1 - \sqrt{\frac{2}{3}}\overline{\alpha}\psi_{\mathrm{i}}\Omega N/M)^{(N-1)/N} & \text{if } 0 \leqslant \overline{N} \leqslant N \neq 0 \end{cases} \tag{42}$$

Recall the hardening law, which relates the image pressure with the state of stress, the state parameter ψ_{i}, and the *deviatoric* component of plastic flow, i.e.

$$\dot{\pi}_{\mathrm{i}} = h(\pi_{\mathrm{i}}^* - \pi_{\mathrm{i}})\dot{\varepsilon}_{\mathrm{s}}^{\mathrm{p}} = \sqrt{\tfrac{2}{3}}h\dot{\lambda}(\pi_{\mathrm{i}}^* - \pi_{\mathrm{i}})\Omega \tag{43}$$

where h is a constant material property, to be calibrated in the finite deformation regime.

Finally, for elastoplasticity the consistency condition necessitates

$$\dot{F} = \mathbf{f} : \dot{\boldsymbol{\tau}} - H\dot{\lambda} = 0, \quad \dot{\lambda} > 0 \tag{44}$$

where $\mathbf{f} := \partial F/\partial\boldsymbol{\tau}$ and H is the so-called hardening modulus defined as

$$H = -\frac{1}{\dot{\lambda}}\frac{\partial F}{\partial \pi_{\mathrm{i}}}\dot{\pi}_{\mathrm{i}} = -\frac{1}{\dot{\lambda}}\left(\frac{p}{\pi_{\mathrm{i}}}\right)^{1/(1-N)} M\dot{\pi}_{\mathrm{i}} \tag{45}$$

Since $p/\pi_{\mathrm{i}} > 0$, the sign of the hardening modulus is governed by the sign of the image pressure rate: $H > 0$ if $\dot{\pi}_{\mathrm{i}} < 0$ (expansion of the yield surface, hardening), $H < 0$ if $\dot{\pi}_{\mathrm{i}} > 0$ (contraction of the yield surface, softening), and $H = 0$ if $\dot{\pi}_{\mathrm{i}} = 0$ (perfect plasticity).

Remark 2

The continuum formulation presented above is a generalization of that presented in Reference [6]. In fact, when $\zeta = \overline{\zeta} = 1 \ \forall\theta \in [0, \pi/3]$, all derivatives of Q with respect to θ drop out, and with $\Omega = \sqrt{\frac{3}{2}}$ we recover the original two-invariant version of the model.

3. NUMERICAL IMPLEMENTATION

We perform a numerical stress point integration based on the product formula algorithm similar to that proposed by Simo [33]. However, here we perform the return mapping in the elastic principal stretch directions to accommodate for nonlinear elasticity. Note that the algorithm provides a closed-form expression for the consistent tangent operator (CTO). The numerical implementation developed below is summarized in Boxes 1 and 2. Box 1 presents a summary of the evolution equations used in the local return mapping, whose recipe is presented in turn in Box 2.

Box 1. Summary of rate equations in three-invariant elastoplastic model for sands.

1. Additive velocity gradient: $\mathbf{l} = \mathbf{l}^{\mathrm{e}} + \mathbf{l}^{\mathrm{p}} \quad \Longrightarrow \quad \mathbf{d} = \mathbf{d}^{\mathrm{e}} + \mathbf{d}^{\mathrm{p}}$
2. Hyperelastic rate equations: $\mathscr{L}_{\mathbf{v}}(\boldsymbol{\tau}) = \mathbf{c}^{\mathrm{e}} : \mathbf{d}^{\mathrm{e}},\ \mathbf{c}^{\mathrm{e}} = 4\mathbf{b}^{\mathrm{e}} \cdot \partial\Psi/(\partial\mathbf{b}^{\mathrm{e}} \otimes \partial\mathbf{b}^{\mathrm{e}}) \cdot \mathbf{b}^{\mathrm{e}}$
3. Nonassociative flow rule: $\mathbf{d}^{\mathrm{p}} = \operatorname{sym} \mathbf{l}^{\mathrm{p}} = \dot{\lambda}\mathbf{q},\ \boldsymbol{\omega}^{\mathrm{p}} = \operatorname{skw} \mathbf{l}^{\mathrm{p}} = \mathbf{0}$
4. State parameter: $\dot{\psi}_{\mathrm{i}} = \dot{v} + \hat{\lambda}\dot{\pi}_{\mathrm{i}}/\pi_{\mathrm{i}}$
5. Hardening law: $\dot{\pi}_{\mathrm{i}} = \sqrt{\frac{2}{3}} h \dot{\lambda}(\pi_{\mathrm{i}}^{*} - \pi_{\mathrm{i}})\Omega$
6. Consistency condition: $\mathbf{f} : \dot{\boldsymbol{\tau}} - H\dot{\lambda} = 0$
7. Kuhn–Tucker optimality conditions: $\dot{\lambda} \geqslant 0,\ F \leqslant 0,\ \dot{\lambda}F = 0$

Box 2. Return mapping algorithm for three-invariant elastoplastic model for sands.

1. Elastic deformation predictor: $\tilde{\mathbf{b}} = \mathbf{f}_{n+1} \cdot \mathbf{b}_{n}^{\mathrm{e}} \cdot \mathbf{f}_{n+1}^{\mathrm{t}}$
2. Elastic stress predictor: $\boldsymbol{\tau}^{\mathrm{tr}} = 2\partial\Psi/\partial\tilde{\mathbf{b}} \cdot \tilde{\mathbf{b}};\ \pi_{\mathrm{i}}^{\mathrm{tr}} = \pi_{\mathrm{i},n}$
3. Check if yielding: $F(\boldsymbol{\tau}^{\mathrm{tr}}, \pi_{\mathrm{i}}^{\mathrm{tr}}) \geqslant 0$?
   No, set $\mathbf{b}^{\mathrm{e}} = \mathbf{b}^{\mathrm{e}\,\mathrm{tr}} \equiv \tilde{\mathbf{b}},\ \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{tr}},\ \pi_{\mathrm{i}} = \pi_{\mathrm{i}}^{\mathrm{tr}}$ and exit
4. Yes, initialize $\Delta\lambda = 0$, build residual $\mathbf{r}(\mathbf{x})$ and iterate for $\mathbf{x}$ (steps 5–8)
5. Spectral decomposition: $\tilde{\mathbf{b}} = \sum_{a=1}^{3} \tilde{\lambda}_a^2 \mathbf{m}^a$
6. Plastic corrector in principal logarithmic stretches: $\varepsilon_a^{\mathrm{e}} = \ln \lambda_a^{\mathrm{e}}$,
   $\tilde{\varepsilon}_a \equiv \varepsilon_a^{\mathrm{e}\,\mathrm{tr}} = \ln \tilde{\lambda}_a,\ \varepsilon_a^{\mathrm{e}} = \varepsilon_a^{\mathrm{e}\,\mathrm{tr}} - \Delta\lambda q_a,\ \tau_a = \partial\overline{\Psi}/\partial\varepsilon_a^{\mathrm{e}}$ for $a = 1, 2, 3$
7. Update plastic internal variable π_{i}:
   (a) Total deformation gradient: $\mathbf{F} = \mathbf{f}_{n+1} \cdot \mathbf{F}_n$
   (b) Specific volume: $v = v_0 \det \mathbf{F} = v_0 J$
   (c) Initialize $\pi_{\mathrm{i}} = \pi_{\mathrm{i},n}$ and iterate for π_{i} (steps 7d–f)
      (d) State parameter: $\psi_{\mathrm{i}} = v - v_{c0} + \hat{\lambda}\ln(-\pi_{\mathrm{i}})$
      (e) Limit hardening plastic variable:
$$\pi_{\mathrm{i}}^{*} = p \times \begin{cases} \exp(\sqrt{\frac{2}{3}}\overline{\alpha}/M\psi_{\mathrm{i}}\Omega) & \text{if } \overline{N} = N = 0 \\ (1 - \sqrt{\frac{2}{3}}\overline{\alpha}\psi_{\mathrm{i}}\Omega N/M)^{(N-1)/N} & \text{if } 0 \leqslant \overline{N} \leqslant N \neq 0 \end{cases}$$
      (f) Plastic internal variable: $\pi_{\mathrm{i}} = \pi_{\mathrm{i},n} + \sqrt{\frac{2}{3}} h \Delta\lambda(\pi_{\mathrm{i}}^{*} - \pi_{\mathrm{i}})\Omega$
8. Discrete consistency condition: $F(p, q, \theta, \pi_{\mathrm{i}}) = 0$
9. Spectral resolution: $\mathbf{b}^{\mathrm{e}} = \sum_{a=1}^{3} \lambda_a^{\mathrm{e}2} \mathbf{m}^a$

3.1. Local return mapping algorithm

Recall the definition of the elastic left Cauchy–Green deformation tensor given in Equation $(2)_2$. From this expression we obtain the trial elastic left Cauchy–Green deformation tensor at time station t_{n+1} by freezing plastic flow i.e.

$$\mathbf{b}^{\mathrm{e\,tr}} \equiv \tilde{\mathbf{b}} = \mathbf{f}_{n+1} \cdot \mathbf{b}_n^{\mathrm{e}} \cdot \mathbf{f}_{n+1}^{\mathrm{t}}, \quad \mathbf{f}_{n+1} = \frac{\partial \mathbf{x}_{n+1}}{\partial \mathbf{x}_n} \tag{46}$$

Using the results presented in Reference [33], the trial elastic deformation tensor can be related to the elastic deformation tensor via the exponential approximation

$$\mathbf{b}^{\mathrm{e}} = \exp(-2\Delta\lambda \mathbf{q}) \cdot \tilde{\mathbf{b}} \tag{47}$$

hence, $\varepsilon_a^{\mathrm{e}} = \varepsilon_a^{\mathrm{e\,tr}} - \Delta\lambda q_a$ and where $\varepsilon_a^{\mathrm{e\,tr}} \equiv \tilde{\varepsilon}_a$ for $a = 1, 2, 3$. From these results in principal logarithmic stretches and utilizing the yield criterion, we formulate the residual vector [6, 15, 34],

$$\mathbf{r}(\mathbf{x}) = \begin{Bmatrix} \varepsilon_1^{\mathrm{e}} - \varepsilon_1^{\mathrm{e\,tr}} + \Delta\lambda q_1 \\ \varepsilon_2^{\mathrm{e}} - \varepsilon_2^{\mathrm{e\,tr}} + \Delta\lambda q_2 \\ \varepsilon_3^{\mathrm{e}} - \varepsilon_3^{\mathrm{e\,tr}} + \Delta\lambda q_3 \\ F \end{Bmatrix} \quad \text{with } \mathbf{x} = \begin{Bmatrix} \varepsilon_1^{\mathrm{e}} \\ \varepsilon_2^{\mathrm{e}} \\ \varepsilon_3^{\mathrm{e}} \\ \Delta\lambda \end{Bmatrix} \tag{48}$$

as the vector of local unknowns. The Newton–Raphson scheme necessitates the Jacobian

$$\mathbf{r}'(\mathbf{x}) = \begin{bmatrix} 1 + \Delta\lambda \partial q_1/\partial \varepsilon_1^{\mathrm{e}} & \Delta\lambda \partial q_1/\partial \varepsilon_2^{\mathrm{e}} & \Delta\lambda \partial q_1/\partial \varepsilon_3^{\mathrm{e}} & q_1 + \Delta\lambda \partial q_1/\partial \Delta\lambda \\ \Delta\lambda \partial q_2/\partial \varepsilon_1^{\mathrm{e}} & 1 + \Delta\lambda \partial q_2/\partial \varepsilon_2^{\mathrm{e}} & \Delta\lambda \partial q_2/\partial \varepsilon_3^{\mathrm{e}} & q_2 + \Delta\lambda \partial q_2/\partial \Delta\lambda \\ \Delta\lambda \partial q_3/\partial \varepsilon_1^{\mathrm{e}} & \Delta\lambda \partial q_3/\partial \varepsilon_2^{\mathrm{e}} & 1 + \Delta\lambda \partial q_3/\partial \varepsilon_3^{\mathrm{e}} & q_3 + \Delta\lambda \partial q_3/\partial \Delta\lambda \\ \partial F/\partial \varepsilon_1^{\mathrm{e}} & \partial F/\partial \varepsilon_2^{\mathrm{e}} & \partial F/\partial \varepsilon_3^{\mathrm{e}} & \partial F/\partial \Delta\lambda \end{bmatrix} \tag{49}$$

The local tangent operator is fully defined once the quantities $\partial q_a/\partial \varepsilon_b^{\mathrm{e}}$, $\partial q_a/\partial \Delta\lambda$, $\partial F/\partial \varepsilon_a^{\mathrm{e}}$ for $a, b = 1, 2, 3$, and $\partial F/\partial \Delta\lambda$ are thoroughly computed. We compute these essential derivatives in what follows.

Let us start by calculating the gradient

$$\frac{\partial q_a}{\partial \varepsilon_b^{\mathrm{e}}} = q_{ac} a_{cb}^{\mathrm{e}} + \frac{\partial q_a}{\partial \pi_{\mathrm{i}}} \frac{\partial \pi_{\mathrm{i}}}{\partial \varepsilon_b^{\mathrm{e}}} \tag{50}$$

where $a_{ab}^{\mathrm{e}} := \partial \tau_a / \partial \varepsilon_a^{\mathrm{e}}$ is the elastic continuum tangent operator in principal directions. Also, we have

$$\frac{\partial q_a}{\partial \pi_{\mathrm{i}}} = \frac{1}{3}\beta \frac{\partial^2 F}{\partial p \partial \pi_{\mathrm{i}}} \delta_a, \quad \frac{\partial^2 F}{\partial p \partial \pi_{\mathrm{i}}} = \frac{1}{1-N} \frac{M}{p} \left(\frac{p}{\pi_{\mathrm{i}}} \right)^{1/(1-N)} \tag{51}$$

The last term of Equation (50) is furnished by the hardening law, which we recall from Equation (43) and then integrate with a Backward Euler scheme to get

$$\pi_{\mathrm{i}} = \pi_{\mathrm{i},n} + \sqrt{\tfrac{2}{3}} h \Delta\lambda (\pi_{\mathrm{i}}^* - \pi_{\mathrm{i}}) \Omega \tag{52}$$

where $\pi_{\mathrm{i},n}$ is the converged value of the image pressure at the previous time step t_n. Therefore,

$$\left(1+\sqrt{\frac{2}{3}}h\Delta\lambda\Omega\right)\frac{\partial\pi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}}=\sqrt{\frac{2}{3}}h\Delta\lambda\left[\left(\frac{\partial\pi_{\mathrm{i}}^*}{\partial\tau_a}a_{ab}^{\mathrm{e}}+\frac{\partial\pi_{\mathrm{i}}^*}{\partial\psi_{\mathrm{i}}}\frac{\partial\psi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}}\right)\Omega+(\pi_{\mathrm{i}}^*-\pi_{\mathrm{i}})\Omega_a a_{ab}^{\mathrm{e}}\right] \tag{53}$$

where we have

$$\frac{\partial\pi_{\mathrm{i}}^*}{\partial\tau_a}=\frac{1}{3}\frac{\pi_{\mathrm{i}}^*}{p}\delta_a+(1-N)\frac{\pi_{\mathrm{i}}^*}{M-\sqrt{2/3}\bar{\alpha}\psi_{\mathrm{i}}\Omega N}\sqrt{\frac{2}{3}}\bar{\alpha}\psi_{\mathrm{i}}\Omega_a \tag{54}$$

and

$$\Omega_a=\frac{1}{\Omega}\left[\frac{3}{2}\bar{\zeta}\,\bar{\zeta}'\theta_a+\bar{\zeta}'q(\bar{\zeta}''\theta_a q+\bar{\zeta}'\sqrt{\frac{3}{2}}\hat{n}_a)\theta_c^2\delta_{\mathrm{c}}+(\bar{\zeta}'q)^2\theta_c\theta_{ca}\right] \tag{55}$$

Finally, recalling the definition of ψ_{i}, we obtain

$$\frac{\partial\pi_{\mathrm{i}}^*}{\partial\psi_{\mathrm{i}}}=\sqrt{\frac{2}{3}}\bar{\alpha}\Omega(1-N)\frac{\pi_{\mathrm{i}}^*}{M-\sqrt{2/3}\bar{\alpha}\psi_{\mathrm{i}}\Omega N}\quad\text{and}\quad\frac{\partial\psi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}}=\frac{\hat{\lambda}}{\pi_{\mathrm{i}}}\frac{\partial\pi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}} \tag{56}$$

Therefore,

$$\frac{\partial\pi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}}=c^{-1}\sqrt{\frac{2}{3}}h\Delta\lambda\left[\Omega\frac{\partial\pi_{\mathrm{i}}^*}{\partial\tau_a}+(\pi_{\mathrm{i}}^*-\pi_{\mathrm{i}})\Omega_a\right]a_{ab}^{\mathrm{e}} \tag{57}$$

where $c=1+\sqrt{\frac{2}{3}}h\Delta\lambda\Omega(1-\hat{\lambda}/\pi_{\mathrm{i}}\partial\pi_{\mathrm{i}}^*/\partial\psi_{\mathrm{i}})$.

It is only left for us to evaluate the derivatives $\partial q_a/\partial\Delta\lambda$, $\partial F/\partial\varepsilon_b^{\mathrm{e}}$, and $\partial F/\partial\Delta\lambda$. By the chain rule, we have

$$\frac{\partial q_a}{\partial\Delta\lambda}=\frac{\partial q_a}{\partial\pi_{\mathrm{i}}}\underbrace{c^{-1}\sqrt{\frac{2}{3}}h\Omega(\pi_{\mathrm{i}}^*-\pi_{\mathrm{i}})}_{\partial\pi_{\mathrm{i}}/\partial\Delta\lambda} \tag{58}$$

Finally, we compute

$$\frac{\partial F}{\partial\varepsilon_b^{\mathrm{e}}}=f_a a_{ab}^{\mathrm{e}}+\frac{\partial F}{\partial\pi_{\mathrm{i}}}\frac{\partial\pi_{\mathrm{i}}}{\partial\varepsilon_b^{\mathrm{e}}}\quad\text{and}\quad\frac{\partial F}{\partial\Delta\lambda}=\frac{\partial F}{\partial\pi_{\mathrm{i}}}\frac{\partial\pi_{\mathrm{i}}}{\partial\Delta\lambda} \tag{59}$$

where

$$f_a\equiv\frac{\partial F}{\partial\tau_a}=\frac{1}{3}\partial_p F\delta_a+\sqrt{\frac{3}{2}}\partial_q F\hat{n}_a+\partial_\theta F\theta_a\quad\text{and}\quad\frac{\partial F}{\partial\pi_{\mathrm{i}}}=M\left(\frac{p}{\pi_{\mathrm{i}}}\right)^{1/(1-N)} \tag{60}$$

In addition to the above local iterative scheme, one sub-local scheme is necessary to solve for π_{i} as it is a nonlinear function of the state of stress, and the state parameter ψ_{i}. This is easily accomplished by introducing the scalar residual

$$r(\pi_{\mathrm{i}})=\pi_{\mathrm{i}}-\pi_{\mathrm{i},n}-\sqrt{\tfrac{2}{3}}h\Delta\lambda(\pi_{\mathrm{i}}^*-\pi_{\mathrm{i}})\Omega \tag{61}$$

with the scalar tangent operator

$$r'(\pi_{\mathrm{i}}) = 1 + \sqrt{\frac{2}{3}} h \Delta\lambda \Omega \left(1 - \frac{\hat{\lambda}}{\pi_{\mathrm{i}}} \frac{\partial \pi_{\mathrm{i}}^*}{\partial \psi_{\mathrm{i}}}\right) \tag{62}$$

Once the sub-local loop is finished and a value for π_{i} is obtained, then ψ_{i} and π_{i}^* can be evaluated and the solution algorithm can proceed to the local Newton–Raphson scheme presented above.

3.2. Consistent tangent in principal directions

Let us start by deriving the elastic tangent in principal directions a_{ab}^{e}. Recall the hyperelastic formulation presented in Section 2.1. We have the strain energy function $\Psi(\varepsilon_{\mathrm{v}}^{\mathrm{e}}, \varepsilon_{\mathrm{s}}^{\mathrm{e}})$ from which the Kirchhoff stress tensor is obtained such that

$$\tau_a = 2 \frac{\partial \Psi}{\partial \lambda_a^{\mathrm{e}2}} \lambda_a^{\mathrm{e}2} = \frac{\partial \overline{\Psi}}{\partial \varepsilon_a^{\mathrm{e}}} = \frac{1}{3} p \delta_a + \sqrt{\frac{2}{3}} q \hat{n}_a \quad \text{(no sum)} \tag{63}$$

where we have used the definitions $p = \partial \Psi / \partial \varepsilon_{\mathrm{v}}^{\mathrm{e}}$ and $q = \partial \Psi / \partial \varepsilon_{\mathrm{s}}^{\mathrm{e}}$. Hence,

$$a_{ab}^{\mathrm{e}} = K^{\mathrm{e}} \delta_a \delta_b + 2\mu^{\mathrm{e}} \left(\delta_{ab} - \tfrac{1}{3} \delta_a \delta_b\right) + \sqrt{\tfrac{2}{3}} d^{\mathrm{e}} (\delta_a \hat{n}_b + \hat{n}_a \delta_b) \tag{64}$$

where $K^{\mathrm{e}} := \partial p / \partial \varepsilon_{\mathrm{v}}^{\mathrm{e}}$ is the elastic bulk modulus, $3\mu^{\mathrm{e}} := \partial q / \partial \varepsilon_{\mathrm{s}}^{\mathrm{e}}$ where μ^{e} is the elastic shear modulus, and $d^{\mathrm{e}} := \partial^2 \Psi / (\partial \varepsilon_{\mathrm{v}}^{\mathrm{e}} \partial \varepsilon_{\mathrm{s}}^{\mathrm{e}})$.

Now, let us define $\tilde{a}_{ab}^{\mathrm{ep}} := \partial \tau_a / \partial \tilde{\varepsilon}_b$ as the *consistent* tangent in principal directions, which we can evaluate with the help of the converged local residual vector via the chain rule. Specifically,

$$\tilde{a}_{ab}^{\mathrm{ep}} = a_{ac}^{\mathrm{e}} a_{cb}^{\mathrm{p}} \quad \text{with } a_{cb}^{\mathrm{p}} \equiv \frac{\partial \varepsilon_c^{\mathrm{e}}}{\partial \tilde{\varepsilon}_b} \tag{65}$$

We can then recall the converged local residual vector $\mathbf{r}(\mathbf{x}) = \mathbf{0}$ and by the chain rule obtain

$$\frac{\partial x_i}{\partial \tilde{\varepsilon}_a} = -b_{ij} \left.\frac{\partial r_j}{\partial \tilde{\varepsilon}_a}\right|_{\mathbf{x}} \tag{66}$$

where $\mathbf{b} := [\mathbf{r}']^{-1}$ and therefore

$$a_{ab}^{\mathrm{p}} = b_{ac} \left(\delta_{cb} - \Delta\lambda \left.\frac{\partial q_c}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}}\right) - b_{a4} \left.\frac{\partial F}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}} \quad \text{for } a, b, c = 1, 2, 3 \tag{67}$$

We also have

$$\left.\frac{\partial q_a}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}} = \frac{\partial q_a}{\partial \pi_{\mathrm{i}}} \left.\frac{\partial \pi_{\mathrm{i}}}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}} \quad \text{and} \quad \left.\frac{\partial F}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}} = \frac{\partial F}{\partial \pi_{\mathrm{i}}} \left.\frac{\partial \pi_{\mathrm{i}}}{\partial \tilde{\varepsilon}_b}\right|_{\mathbf{x}} \tag{68}$$

where

$$\left.\frac{\partial \pi_{\mathrm{i}}}{\partial \tilde{\varepsilon}_a}\right|_{\mathbf{x}} = c^{-1} \sqrt{\frac{2}{3}} h \Delta\lambda \Omega v \frac{\partial \pi_{\mathrm{i}}^*}{\partial \psi_{\mathrm{i}}} \delta_a \tag{69}$$

We note that in the special case of pure elasticity, the tangent $a^{\mathrm{p}}_{ab} \equiv \delta_{ab}$ and thus, $\tilde{a}^{\mathrm{ep}}_{ab} \equiv a^{\mathrm{e}}_{ab}$. It will be shown in the next section, that the consistent tangent in principal directions contributes to the global consistent tangent operator, which for this particular class of return mapping algorithm can be obtained in closed-form.

4. CONSISTENT TANGENT OPERATORS IN SPECTRAL FORM BASED ON LIE DERIVATIVES

It is generally recognized that a consistent linearization of the so-called weak form of balance of linear momentum is critical for optimal performance of an iterative algorithm [20, 35, 36]. In particular, in the presence of nonlinear kinematics the concept of directional derivatives plays a key role in the linearization process. We proceed to developing the CTO for the case of elastoplasticity. For the elastoplastic formulation, isotropy (i.e. $\Psi = \Psi(\lambda^{\mathrm{e}}_1, \lambda^{\mathrm{e}}_2, \lambda^{\mathrm{e}}_3)$) furnishes a connection between the elastic and trial kinematical quantities and consequently, it is a crucial component in our developments. The result hinges on two main results.

Proposition 1
The Kirchhoff stress tensor can be obtained using the trial elastic left Cauchy–Green deformation tensor $\mathbf{b}^{\mathrm{e\,tr}} \equiv \tilde{\mathbf{b}} := \tilde{\mathbf{F}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}}$, i.e.

$$\boldsymbol{\tau} = 2\frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{b}}} \cdot \tilde{\mathbf{b}} \tag{70}$$

where $\mathbf{F}^{\mathrm{e\,tr}} \equiv \tilde{\mathbf{F}} = \sum_{a=1}^{3} \tilde{\lambda}_a \mathbf{n}^a \otimes \mathbf{N}^a$ is the trial elastic deformation gradient.

Proof
By definition,

$$\boldsymbol{\tau} = 2\frac{\partial \Psi(\mathbf{b}^{\mathrm{e}})}{\partial \mathbf{b}^{\mathrm{e}}} \cdot \mathbf{b}^{\mathrm{e}} \tag{71}$$

We also know that we can decompose $\tilde{\mathbf{b}}$ and $\mathbf{b}^{\mathrm{e}}$ spectrally and obtain, respectively

$$\tilde{\mathbf{b}} = \sum_{a=1}^{3} \tilde{\lambda}^2_a \mathbf{m}^a \quad \text{and} \quad \mathbf{b}^{\mathrm{e}} = \sum_{a=1}^{3} \lambda^{\mathrm{e}2}_a \mathbf{m}^a \tag{72}$$

where we note the fact that the tensors $\tilde{\mathbf{b}}$ and $\mathbf{b}^{\mathrm{e}}$ have the same eigenvectors, a by-product of the return mapping algorithm. Then we use Equations (71) and (72) along with the chain rule and write

$$\boldsymbol{\tau} = 2\sum_{a=1}^{3} \frac{\partial \tilde{\Psi}}{\partial \tilde{\lambda}^2_a} \frac{\partial \tilde{\lambda}^2_a}{\partial \lambda^{\mathrm{e}2}_a} \lambda^{\mathrm{e}2}_a \mathbf{m}^a = 2\sum_{a=1}^{3} \frac{\partial \tilde{\Psi}}{\partial \tilde{\lambda}^2_a} \frac{\tilde{\lambda}^2_a}{\lambda^{\mathrm{e}2}_a} \lambda^{\mathrm{e}2}_a \mathbf{m}^a = 2\frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{b}}} \cdot \tilde{\mathbf{b}} \tag{73}$$

where we have exploited the algorithmic relationship between λ^{e}_a and $\tilde{\lambda}_a$ (cf. Equation (4.5,b) in Reference [33]) and the isotropy of the strain-energy function. □

Proposition 2
Define the trial elastic right Cauchy–Green deformation tensor $\tilde{\mathbf{C}} := \tilde{\mathbf{F}}^{\mathrm{t}} \cdot \tilde{\mathbf{F}} = \sum_{a=1}^{3} \tilde{\lambda}_a^2 \mathbf{M}^a$, then

$$\boldsymbol{\tau} = 2 \frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{b}}} \cdot \tilde{\mathbf{b}} = 2\tilde{\mathbf{F}} \cdot \underbrace{\frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{C}}}}_{:=\frac{1}{2}\tilde{\mathbf{S}}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}} = \tilde{\mathbf{F}} \cdot \tilde{\mathbf{S}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}} \tag{74}$$

Proof
This is a standard result in continuum mechanics, but we will prove it here for completeness. We start by taking the time derivative of the strain-energy function

$$\begin{aligned} \dot{\tilde{\Psi}} &= \frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{b}}} : \dot{\tilde{\mathbf{b}}} = \frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{C}}} : \dot{\tilde{\mathbf{C}}} \\ &= 2 \left[\frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{b}}} \cdot \tilde{\mathbf{F}} \right] : \dot{\tilde{\mathbf{F}}} = 2 \left[\tilde{\mathbf{F}} \cdot \frac{\partial \tilde{\Psi}}{\partial \tilde{\mathbf{C}}} \right] : \dot{\tilde{\mathbf{F}}} \end{aligned} \tag{75}$$

where we have exploited the symmetries of $\tilde{\mathbf{b}}$ and $\tilde{\mathbf{C}}$, the isotropy of $\tilde{\Psi}$, and the chain rule. Consequently, since the expression above must hold for *all* $\tilde{\mathbf{F}}$, we conclude $\partial \tilde{\Psi} / \partial \tilde{\mathbf{b}} = \tilde{\mathbf{F}} \cdot \partial \tilde{\Psi} / \partial \tilde{\mathbf{C}} \cdot \tilde{\mathbf{F}}^{-1}$, from where Equation (74) follows. □

With Equations (70) and (74), it is then easy to obtain a closed-form expression for the CTO by utilizing the spectral decomposition approach along with the relationship between the material time derivative and the Lie derivative of the Kirchhoff stress tensor. We start by taking the material time derivative of the Kirchhoff stress tensor

$$\begin{aligned} \dot{\boldsymbol{\tau}} &= \frac{\mathrm{d}}{\mathrm{d}t} (\tilde{\mathbf{F}} \cdot \tilde{\mathbf{S}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}}) \\ \dot{\boldsymbol{\tau}} &= \dot{\tilde{\mathbf{F}}} \cdot \tilde{\mathbf{S}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}} + \tilde{\mathbf{F}} \cdot \tilde{\mathbf{S}} \cdot \dot{\tilde{\mathbf{F}}}^{\mathrm{t}} + \tilde{\mathbf{F}} \cdot \underbrace{\frac{\partial \tilde{\mathbf{S}}}{\partial \tilde{\mathbf{C}}}}_{:=\tilde{\mathbf{A}}} : \dot{\tilde{\mathbf{C}}} \cdot \tilde{\mathbf{F}}^{\mathrm{t}} \end{aligned} \tag{76}$$

It is then necessary to explicate all three terms above. We recall $\tilde{\mathbf{F}} = \mathbf{F} \cdot \mathbf{F}_n^{\mathrm{p}-1}$ and thus $\dot{\tilde{\mathbf{F}}} = \dot{\mathbf{F}} \cdot \mathbf{F}_n^{\mathrm{p}-1} = \mathbf{l} \cdot \tilde{\mathbf{F}}$, where $\mathbf{F}_n^{\mathrm{p}}$ is the plastic deformation gradient at time t_n. Using this result, we can calculate

$$\dot{\tilde{\mathbf{C}}} = \tilde{\mathbf{F}}^{\mathrm{t}} \cdot \mathbf{l} \cdot \tilde{\mathbf{F}} + \tilde{\mathbf{F}}^{\mathrm{t}} \cdot \mathbf{l}^{\mathrm{t}} \cdot \tilde{\mathbf{F}} = 2\tilde{\mathbf{F}}^{\mathrm{t}} \cdot \mathbf{d} \cdot \tilde{\mathbf{F}} \tag{77}$$

It is now possible to rewrite Equation (76) as

$$\dot{\boldsymbol{\tau}} = \mathbf{l} \cdot \boldsymbol{\tau} + \boldsymbol{\tau} \cdot \mathbf{l}^{\mathrm{t}} + \tilde{\mathbf{c}} : \mathbf{d} \tag{78}$$

where we identify $\mathscr{L}_{\mathbf{v}}(\boldsymbol{\tau}) = \tilde{\mathbf{c}} : \mathbf{d}$ and

$$\tilde{c}_{ijkl} = 2\tilde{F}_{iI}\tilde{F}_{jJ}\tilde{F}_{kK}\tilde{F}_{lL}\tilde{A}_{IJKL} \tag{79}$$

It is clear that since $\tilde{\Psi}=\tilde{\Psi}(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3)$, then it follows that $\tilde{\mathbf{S}}=\sum_{a=1}^{3} \tilde{S}_a \mathbf{M}^a$ and $\tilde{\mathbf{C}}=\sum_{a=1}^{3} \tilde{\lambda}_a^2 \mathbf{M}^a$ are coaxial and consequently

$$\tilde{\mathbf{A}} = \sum_{a=1}^{3}\sum_{b=1}^{3} \frac{\partial \tilde{S}_a}{\partial \tilde{\lambda}_b^2} \mathbf{M}^a \otimes \mathbf{M}^b + \frac{1}{2}\sum_{a=1}^{3}\sum_{b\neq a} \frac{\tilde{S}_b - \tilde{S}_a}{\tilde{\lambda}_b^2 - \tilde{\lambda}_a^2} (\mathbf{M}^{ab} \otimes \mathbf{M}^{ab} + \mathbf{M}^{ab} \otimes \mathbf{M}^{ba}) \tag{80}$$

and therefore, by pushing $\tilde{\mathbf{A}}$ forward with Equation (79) and using the fact that $\tilde{\lambda}_a^2 \tilde{S}_a = \tau_a$ we obtain

$$\tilde{\mathbf{c}} = \sum_{a=1}^{3}\sum_{b=1}^{3} \underbrace{(\tilde{a}_{ab}^{\mathrm{ep}} - 2\tau_a \delta_{ab})}_{:=\tilde{c}_{ab}} \mathbf{m}^a \otimes \mathbf{m}^b + \sum_{a=1}^{3}\sum_{b\neq a} \underbrace{\frac{\tau_b \tilde{\lambda}_a^2 - \tau_a \tilde{\lambda}_b^2}{\tilde{\lambda}_b^2 - \tilde{\lambda}_a^2}}_{:=\tilde{\gamma}_{ab}} (\mathbf{m}^{ab} \otimes \mathbf{m}^{ab} + \mathbf{m}^{ab} \otimes \mathbf{m}^{ba}) \tag{81}$$

where $\tilde{a}_{ab}^{\mathrm{ep}} := \partial \tau_a / \partial \tilde{\varepsilon}_b$, $\tilde{\varepsilon}_b := \ln \tilde{\lambda}_b$, and where $\mathbf{M}^a := \mathbf{N}^a \otimes \mathbf{N}^a$, $\mathbf{M}^{ab} := \mathbf{N}^a \otimes \mathbf{N}^b$ and $\mathbf{m}^{ab} := \mathbf{n}^a \otimes \mathbf{n}^b$. We observe from Equation (81) that the CTO possesses minor symmetries but lacks major symmetry if $\tilde{a}_{ab}^{\mathrm{ep}} \neq \tilde{a}_{ba}^{\mathrm{ep}}$, which is certainly the case for nonassociative plasticity or when the integration algorithm destroys the symmetry of the tangent operator [12, 37].

5. SEARCH ALGORITHM IN PRINCIPAL STRESS SPACE

Consider the expression for the total tangent operator i.e. $\boldsymbol{a}^{\mathrm{ep}} = \mathbf{c}^{\mathrm{ep}} + \boldsymbol{\tau} \oplus \mathbf{1}$, where the elasto-plastic tangent operator is form-identical to the CTO derived above, namely,

$$\mathbf{c}^{\mathrm{ep}} = \sum_{a=1}^{3}\sum_{b=1}^{3} c_{ab}^{\mathrm{ep}} \mathbf{m}^a \otimes \mathbf{m}^b + \sum_{a=1}^{3}\sum_{b\neq a} \gamma_{ab} (\mathbf{m}^{ab} \otimes \mathbf{m}^{ab} + \mathbf{m}^{ab} \otimes \mathbf{m}^{ba}) \tag{82}$$

We use this operator in defining the Eulerian acoustic tensor for nonlinear kinematics [18]

$$A_{ik} \equiv n_j a_{ijkl}^{\mathrm{ep}} n_l \tag{83}$$

where n_i for $i = 1, 2, 3$ are the components of the unit vector $\mathbf{n}$ normal to an impending shear band in the current configuration. Since the principal directions of the stress tensor $\boldsymbol{\tau}$ span $\mathbb{R}^3$, we can construct *any* vector as a linear combination of these eigenvectors. Specifically,

$$\mathbf{n} = \sum_{a=1}^{3} \alpha_a \mathbf{n}^a \tag{84}$$

which we can then combine with (83) and (82) to get the spectral components of the acoustic tensor

$$\hat{A}_{ab} = \begin{cases} \alpha_a^2 c_{aa}^{\mathrm{ep}} + \sigma + \sum_{c\neq a} \alpha_c^2 \gamma_{ca} & \text{if } a = b \\ \alpha_a (c_{ab}^{\mathrm{ep}} + \gamma_{ab}) \alpha_b & \text{if } a \neq b \end{cases} \quad \text{(no sum)} \tag{85}$$

where $\sigma := \mathbf{n} \cdot \boldsymbol{\tau} \cdot \mathbf{n}$ is the component of the stress tensor normal to the shear band and $\hat{A}_{ab} := \mathbf{n}^a \cdot \mathbf{A} \cdot \mathbf{n}^b$ are the components of the acoustic tensor in principal direction basis.

We recall that a necessary condition for localization is [6, 18, 38],

$$\mathscr{F}(\mathbf{A}) = \inf|_{\mathbf{n}} \det(\mathbf{A}) = 0 \tag{86}$$

Note that the determinant of a second-order tensor is an invariant quantity under rotations, i.e. $\det(\mathbf{A}) = \det(\hat{\mathbf{A}})$, so this allows us to use the spectral form of the acoustic tensor only. Defining $f = \det(\hat{\mathbf{A}})$, the localization condition can be rewritten in the equivalent form

$$\frac{\partial f}{\partial \alpha_a} = 0, \quad a = 1, 2, 3 \tag{87}$$

which provides three optimality conditions that need to be satisfied in order to obtain a local minimum. We can further reduce the number of unknowns by introducing the spherical coordinates $(\theta, \phi, \|\mathbf{n}\|_2)$ in principal space (see Figure 2). We require $\|\mathbf{n}\|_2 \equiv 1$, which introduces a trivial constraint on the system. Some algorithms use this constraint in the framework of Lagrange multipliers to solve for $\mathscr{F}$ (e.g. see Reference [39]).

From Figure 2, we have the coordinate relations

$$\alpha_1 = \sin\theta \sin\phi, \quad \alpha_2 = \cos\phi, \quad \alpha_3 = \cos\theta \sin\phi \tag{88}$$

We can then use these relations to rewrite the localization condition by defining the vector $\boldsymbol{\beta} = \{\theta, \phi\}^{\mathrm{t}}$, and by the chain rule,

$$\frac{\partial f}{\partial \beta_i} = \sum_{a=1}^{3} \frac{\partial f}{\partial \alpha_a} \frac{\partial \alpha_a}{\partial \beta_i} = 0, \quad i = 1, 2 \tag{89}$$

Due to the highly nonlinear nature of the optimality conditions in Equation (89), we resort to a Newton–Raphson iterative procedure by expanding the residual around a previous iteration

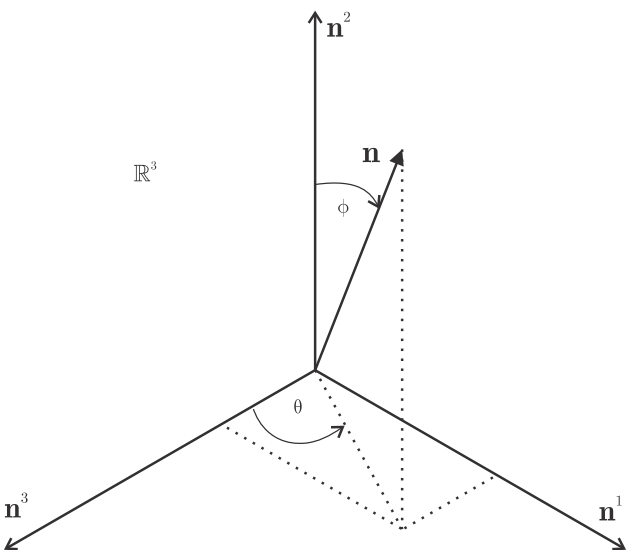

Figure 2. Reference frame in principal directions basis.

and equating it to zero i.e.

$$\left.\frac{\partial f}{\partial \boldsymbol{\beta}}\right|_{k+1} \approx \left.\frac{\partial f}{\partial \boldsymbol{\beta}}\right|_{k} + \left.\frac{\partial^2 f}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}}\right|_{k} \cdot \underbrace{(\boldsymbol{\beta}_{k+1} - \boldsymbol{\beta}_k)}_{:=\Delta\boldsymbol{\beta}} = 0 \tag{90}$$

and hence

$$\Delta\boldsymbol{\beta} = -\left(\frac{\partial^2 f}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}}\right)^{-1}_{k} \cdot \left.\frac{\partial f}{\partial \boldsymbol{\beta}}\right|_{k} \tag{91}$$

The first approximation $\boldsymbol{\beta}_0$ is furnished by a coarse sweep over half a unit ball, i.e. $[0, \pi] \times [0, \pi]$, where the global minimum of the function f is attained.

Now, it remains to explicate the right-hand side of Equation (91). Using the fact that $\partial \det(\hat{\mathbf{A}})/\partial \hat{\mathbf{A}} = \det(\hat{\mathbf{A}})\hat{\mathbf{A}}^{-\mathrm{t}}$, along with the chain rule, we get

$$\frac{\partial f}{\partial \boldsymbol{\beta}} = f\hat{\mathbf{A}}^{-\mathrm{t}} : \frac{\partial \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}} \cdot \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\beta}} \tag{92}$$

where $\partial\hat{\mathbf{A}}/\partial\boldsymbol{\alpha}$ is a third-order tensor with components (cf. Equation (85))

$$\frac{\partial \hat{A}_{ab}}{\partial \alpha_c} = \begin{cases} 2\alpha_a \delta_{ac} c^{\mathrm{ep}}_{aa} + 2\alpha_c \tau_c + \underbrace{2\alpha_c \gamma_{ca}}_{c\neq a} & \text{if } a = b \\ & \text{(no sum)} \\ \delta_{ac}(c^{\mathrm{ep}}_{ab} + \gamma_{ab})\alpha_b + \alpha_a(c^{\mathrm{ep}}_{ab} + \gamma_{ab})\delta_{bc} & \text{if } a \neq b \end{cases} \tag{93}$$

with δ_{ab} being the Kronecker delta. Further, the components of the matrix $\partial\boldsymbol{\alpha}/\partial\boldsymbol{\beta}$ can be easily computed to yield

$$\frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} \partial_\theta \alpha_1 & \partial_\phi \alpha_1 \\ \partial_\theta \alpha_2 & \partial_\phi \alpha_2 \\ \partial_\theta \alpha_3 & \partial_\phi \alpha_3 \end{bmatrix} \tag{94}$$

The Jacobian matrix can be calculated with the aid of the chain rule as

$$\frac{\partial^2 f}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}} = \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\beta}} \cdot \mathbf{J} \cdot \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\beta}} + \frac{\partial f}{\partial \boldsymbol{\alpha}} \cdot \frac{\partial^2 \alpha}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}} \tag{95}$$

where

$$\mathbf{J} = f\left[\left(\hat{\mathbf{A}}^{-\mathrm{t}} : \frac{\partial \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}}\right) \otimes \left(\hat{\mathbf{A}}^{-\mathrm{t}} : \frac{\partial \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}}\right) - \frac{\partial \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}} : \hat{\mathbf{A}}^{-\mathrm{t}} \ominus \hat{\mathbf{A}}^{-1} : \frac{\partial \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}} + \hat{\mathbf{A}}^{-\mathrm{t}} : \frac{\partial^2 \hat{\mathbf{A}}}{\partial \boldsymbol{\alpha}\partial \boldsymbol{\alpha}}\right] \tag{96}$$

Finally, the fourth-order tensor $\partial^2\hat{\mathbf{A}}/(\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha})$ can be written in component form as

$$\frac{1}{2}\frac{\partial \hat{A}_{ab}}{\partial \alpha_c \partial \alpha_d} = \begin{cases} \delta_{ac}\delta_{bd} c^{\mathrm{ep}}_{aa} + \delta_{cd}\tau_c + \underbrace{\delta_{cd}\gamma_{ca}}_{c\neq a} & \text{if } a = b \\ & \text{(no sum)} \\ (c^{\mathrm{ep}}_{ab} + \gamma_{ab}).\mathscr{I}_{abcd} & \text{if } a \neq b \end{cases} \tag{97}$$

where $\mathscr{I}_{abcd} = \frac{1}{2}(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc})$ is the fourth-order identity tensor.

Note that the tensors presented in Equations (93) and (97) are very sparse, thus making the assembly of the residual vector and the Jacobian matrix very straightforward. Also, the Hessian matrix in (97) is only calculated once in the algorithm as it is not a function of the α_a's. At this point, it is easy to recognize the importance of the fact that since the algorithm is based on Newton–Raphson scheme, asymptotic quadratic rate of convergence is achieved and as a result the algorithm only needs to iterate a few times to find the global minimum, provided, of course, that the first guess is close enough to the solution.

It has been shown by several researchers that for a wide class of elastoplastic models, at least one of the α_a's is zero and therefore the search only needs to take place on the principal planes [38, 40–42]. For example, classical models such as von-Misses, Drucker–Prager, and Mohr–Coulomb can all be shown to localize within one of the principal planes, even for the case of finite strains, provided a suitable form of the elastic continuum tangent is utilized (isotropic). Also, the infinitesimal version of the elastoplastic model presented herein can be shown to belong to the class of constitutive models described above. Due to the spectral nature of the algorithm presented here, the on-plane feature is easily activated, thus making the search very efficient. This feature is not present in some of the previously proposed search algorithms available in the literature, see for example the works by Ortiz *et al*. [39] and Mosler [43], which always perform fully three-dimensional searches. Simulations underscoring the main features of the algorithm such as asymptotic quadratic rate and the on-plane feature are included in Section 6.

Remark 3

Note that the formulation presented above is general and applicable to models with linear and nonlinear kinematics. The same algorithm can be easily utilized for the case of infinitesimal deformations provided that the stress terms are set identically equal to zero, for example.

Remark 4

As pointed out first in Reference [16], the *consistent* tangent operator can be used in lieu of its continuum counterpart to search for the necessary condition for localization, provided a small enough load step is taken. This can be accomplished, in light of the above-described search algorithm, by replacing c_{ab}^{ep} with $\tilde{c}_{ab}$ and γ_{ab} with $\tilde{\gamma}_{ab}$. It is very convenient to use the algorithmic operator as it is already available from the material subroutine and therefore the continuum tangent is not to be used explicitly.

6. NUMERICAL EXAMPLES

In this section, we conduct local stress-point simulations to illustrate the main features of the three-invariant constitutive formulation developed in this work. Localization analyses are conducted at this level and the convergence profile of the search algorithm for the minimum determinant is reported. Then, a series of boundary-value problems at the finite element level are performed with the objective of studying the behaviour of granular bodies exhibiting unstructured random density fields. The simulations focus on *detecting* bifurcation and no effort is made to try to capture post-bifurcation behaviour. Hence all simulations are stopped once bifurcation is detected.

6.1. Stress-point simulations

At the stress-point level, we performed two strain-driven simulations to highlight the difference in the responses resulting from the inclusion of the third stress invariant and to test the performance of the search algorithm for the minimum value of the localization function. Two material points with identical properties, except for the ellipticity parameters ρ and $\overline{\rho}$, were strained until they reached localization. The hyperelastic parameters are as follows: compressibility coefficient $\hat{\kappa} = 0.01$; reference elastic volumetric strain $\varepsilon_{v0}^{e} = 0$ at a reference pressure $p_0 = -100\,\text{kPa}$; initial shear modulus $\mu_0 = 5400\,\text{kPa}$; and coupling constant $\alpha_0 = 0$ (see Section 2.1 for notation). The plasticity parameters are: plastic compressibility coefficient $\hat{\lambda} = 0.0135$; critical state parameter $M = 1.2$; shape parameters $N = 0.4$ and $\overline{N} = 0.2$ (nonassociative volumetric plastic flow); and hardening constant $h = 280$ (see Section 2.2 for notation). Both material points were initially denser than critical with an initial specific volume $v = 1.59$ and a reference specific volume $v_{c0} \approx 1.81$. The only difference in material properties is the ellipticity: one material point had $\rho = \overline{\rho} = 1$ (two-invariant formulation, circular cross-section and associative flow on the deviatoric plane) and the other had $\rho = 0.7$ and $\overline{\rho} = 0.8$ (noncircular, convex cross-section, nonassociative flow on deviatoric plane). Both material points were modelled using the Willam–Warnke shape function in Equation (12).

The loading protocol is as follows. We prescribed two equal sets of relative deformation gradient $\mathbf{f}_{n+1} = \partial \mathbf{x}_{n+1} / \partial \mathbf{x}_n$: the first set was prescribed over n_1 steps, followed by the second for n_2 steps until localization was reached. The two relative deformation gradients are

$$\mathbf{f}_1 = \begin{bmatrix} 1+\lambda_2 & 0 & 0 \\ 0 & 1-\lambda_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{f}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-\lambda_2 & 0 \\ 0 & 0 & 1+\lambda_1 \end{bmatrix} \tag{98}$$

hence, the total deformation gradient is

$$\mathbf{F} = \mathbf{f}_2^{n_2} \cdot \mathbf{f}_1^{n_1} \tag{99}$$

where $\lambda_1 = 1 \times 10^{-3}$ and $\lambda_2 = 4 \times 10^{-4}$. The material points were mostly compressed in the first $n_1 = 10$ steps, and then mostly stretched in the subsequent n_2 steps, hence the total number of steps is given by $n = n_1 + n_2$. The stress paths followed by the two material points are shown in Figures 3 and 4, with Figure 3 showing the stress paths on a meridian plane and Figure 4 showing them on a deviatoric plane. The paths are plotted using rotated principal stress axis as described in Reference [15]. The trajectories followed by the two stress points are similar. The stress paths started at point O and then were loaded to point A' after the first n_1 steps, and subsequently deformed to points B' and C', where they localized. Similarly, the yield surfaces expanded from A to B to C, without softening.

The sample with $\rho = 0.7$ localized first at $n = 22$, whereas the sample with $\rho = 1$ localized at $n = 26$. The function $\mathscr{F}(\mathbf{A})$ for both stress points is plotted in Figure 5 where we see minor differences in the trajectories, yet the stress point with $\rho = 0.7$ localized sooner. Hence, the third invariant enhances strain localization. The values for the function $\mathscr{F}(\mathbf{A})$ are obtained by the search algorithm outlined in Section 5 by sweeping the $\mathbb{R}^3$ space via the angles θ and ϕ.

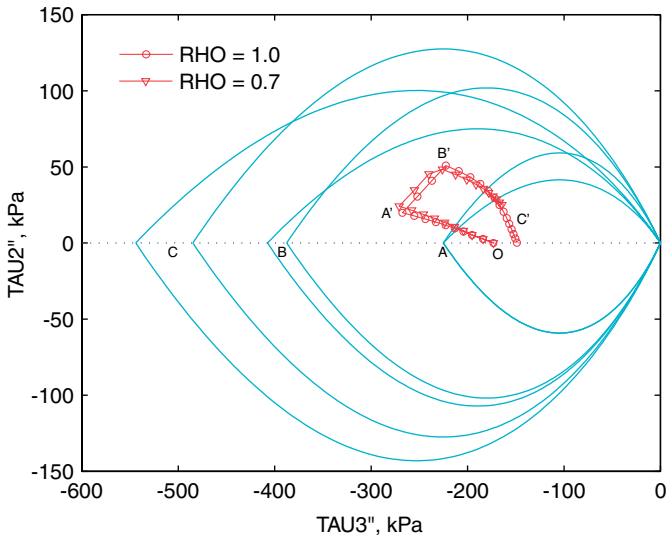

Figure 3. Stress paths on meridian plane. Yield surfaces expand from A to B and to C whereas stress paths follow O–A'–B'–C' trajectory.

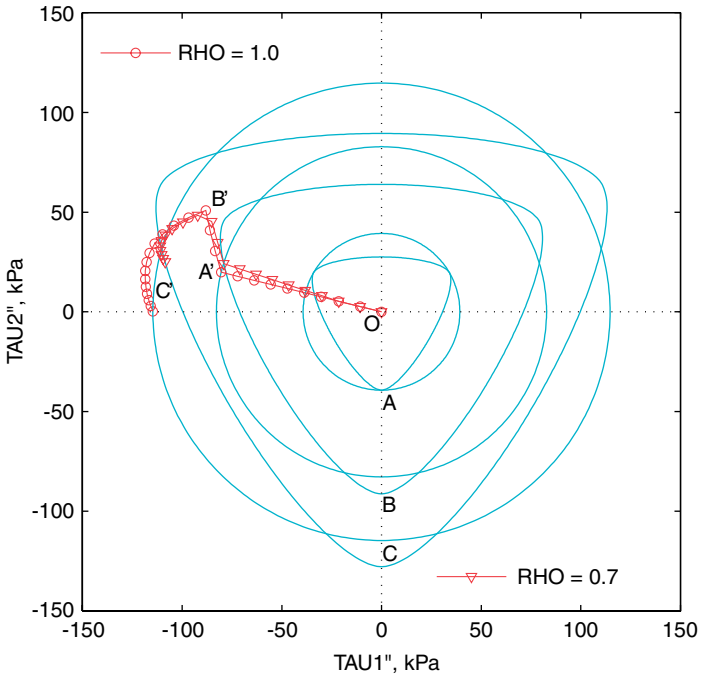

Figure 4. Stress paths on deviatoric plane. Yield surfaces expand from A to B and to C whereas stress paths follow O–A'–B'–C' trajectory.

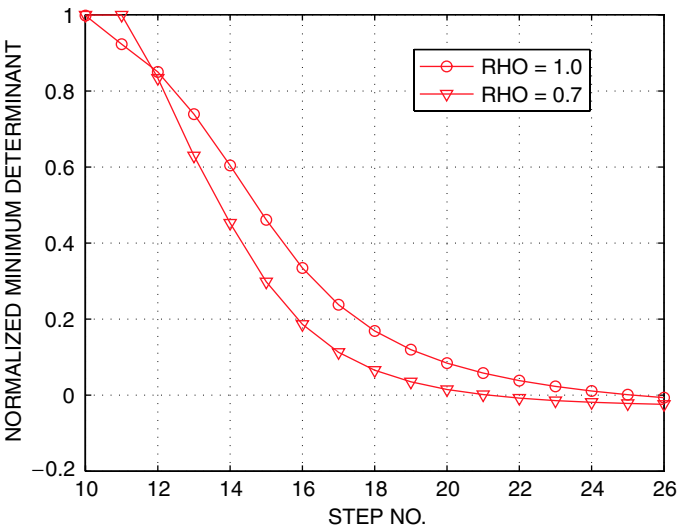

Figure 5. Minimum determinant of the acoustic tensor at various load steps.

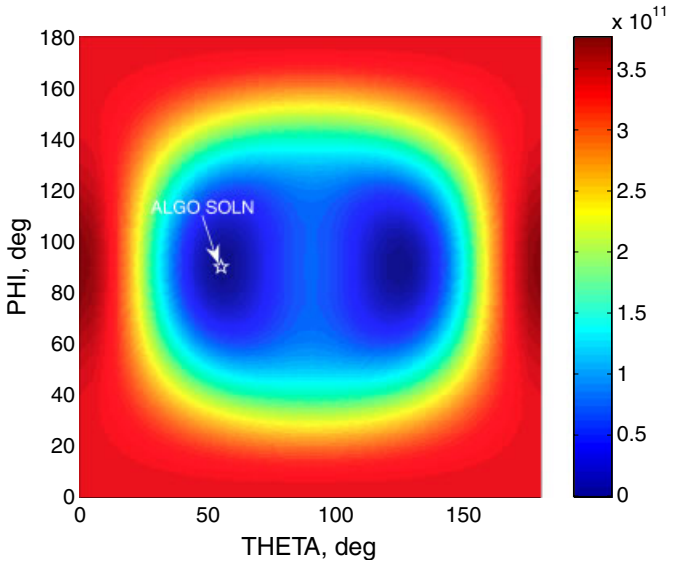

Figure 6. Profile of the determinant of the acoustic tensor for three-invariant model with $\rho = 0.7$ at onset of localization.

One such a profile swept by the search algorithm is shown in Figure 6, where the determinant of the acoustic tensor is plotted as a function of θ and ϕ at the onset of localization for the material point with $\rho = 0.7$. One such profile is swept at each time step where the search

algorithm finds one of usually two wells and then returns a solution. One interesting note from Figure 6 is that localization was achieved at $\phi = \pi/2$ (see end of Section 5 for discussion). Finally, the convergence profile for the search algorithm at various load steps is reported in Figure 7 where optimal quadratic convergence is observed.

6.2. *Simulations with cubical specimens*

Randomization of the density field in a specimen is achieved through the void ratio e, defined as the ratio between the volume occupied by empty voids and the intrinsic volume of the solid phase, i.e. $e := V_v / V_s$. The main issue concerns finding a plausible probability density function describing the natural dispersion of voids in a sample of granular materials. Shahinpoor [44] used statistical mechanical theory on a collection of Voronoi cells to obtain an exact expression for the probability density function for the distributions of void ratio in a random aggregate of granular material. The probability distribution function for e was established to be a truncated exponential distribution. It also has been used by Nübel and Huang [45] to study localized deformation patterns in granular media.

Here, we adopted a truncated exponential density function for e of the form

$$f_e(e) = \frac{\gamma \exp(-\gamma e)}{\exp(-\gamma e_d) - \exp(-\gamma e_l)} \tag{100}$$

where e_d is the lower bound of the distribution corresponding to the densest state, e_l is the upper bound of the distribution corresponding to the loosest state, and γ plays the role of a fitting parameter. We used these bounds to restrict the amount of dispersion in the density of the

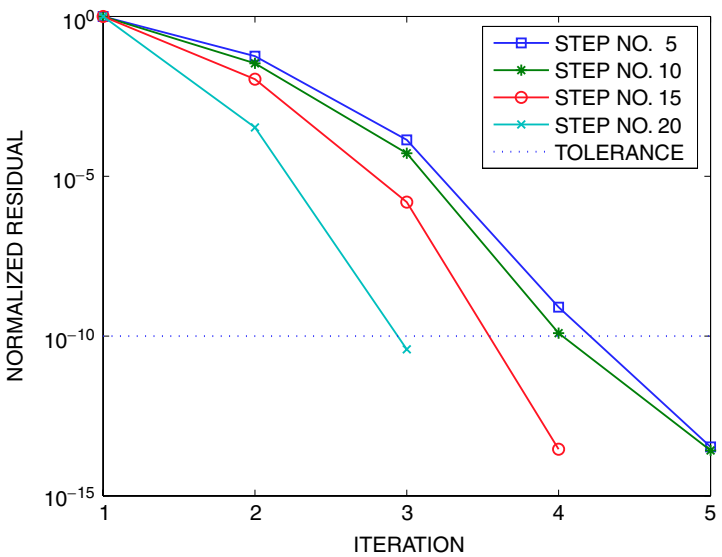

Figure 7. Convergence profile for search algorithm at various load steps.

sample, thus controlling the amount of inhomogeneities. The expected value of the distribution is given by [46]

$$\overline{e} = \int_{e_d}^{e_l} e f_e(e)\, \mathrm{d}e \tag{101}$$

where

$$\overline{e} = \frac{1}{\gamma} + \frac{e_d \exp(-\gamma e_d) - e_l \exp(-\gamma e_l)}{\exp(-\gamma e_d) - \exp(-\gamma e_l)} \tag{102}$$

Hence, we can *prescribe* the expected value or mean void ratio $\overline{e}$ subject to the upper and lower bounds e_d and e_l, respectively, provided we solve for the fitting coefficient γ. This is easily accomplished by a Newton–Raphson scheme on the residual

$$r(\gamma) = \overline{e} - \frac{1}{\gamma} - \frac{e_d \exp(-\gamma e_d) - e_l \exp(-\gamma e_l)}{\exp(-\gamma e_d) - \exp(-\gamma e_l)} \tag{103}$$

The random density field was then generated on the assumption that the values of void ratio in space are mutually independent. For this reason, these random fields are called 'unstructured' due to the lack of correlation in the density field between points A and B in the same sample.

We then sheared in compression two rectangular specimens with dimensions $1 \times 1 \times 2$ m and assumed to behave according to the constitutive framework presented above, up to the onset of localization. The material properties for both specimens were identical to those of the stress-point with $\rho = \overline{\rho} = 0.78$ and $v_{c0} \approx 1.85$ discussed in the previous section. The Argyris–Gudehus shape function given in Equation (11) was utilized in the simulations. At the structural level the density field was generated using the exponential distribution described above with a mean void ratio $\overline{e} = 0.63$ and lower and upper bounds $e_d = 0.54$ and $e_l = 0.64$, respectively. Hence, the two specimens, which we will call 'INHOMOGENEOUS 1' and 'INHOMOGENEOUS 2' for identification purposes, represent two different realizations of the same distribution function for the void ratio field, with otherwise identical mechanical properties. Figure 8 shows the discretized specimens with their respective initial specific volume field $v := 1+e$. The rectangular domains were discretized using 2000 trilinear hexahedral elements equipped with the B-bar method for nonlinear kinematics using the so-called mean dilation technique [47, 48] (see Reference [49] for a survey on the B-bar method).

The boundary conditions applied to the specimens were as follows. All four lateral faces were initially subjected to a constant confining pressure of 100 kPa (Newman BCs). The bottom and top faces of the specimens ($z = 0, 1$ m) were supported by rollers with a pin at the $(0, 0, 0)$ m point for stability (Dirichlet BCs). The bottom face was constrained from displacement in the z direction, whereas the top face was subjected to a vertical displacement responsible for shearing the samples in compression. The samples were loaded in two phases. Phase one, an all-around confining pressure was applied to consolidate the sample followed by phase two, where the top face was displaced vertically. The objective of this loading protocol, favouring homogeneous deformations in perfectly homogeneous samples, is twofold: to observe if any significant differences are introduced by perturbing the density field by comparing the homogeneous response against perturbed responses from the samples shown in Figure 8, and to compare the responses of two inhomogeneous samples from two realizations of the density field. To this end, a homogeneous sample with constant void ratio $v = 1.63$ was subjected to the same BCs described above, and its response compared against its inhomogeneous counterparts.

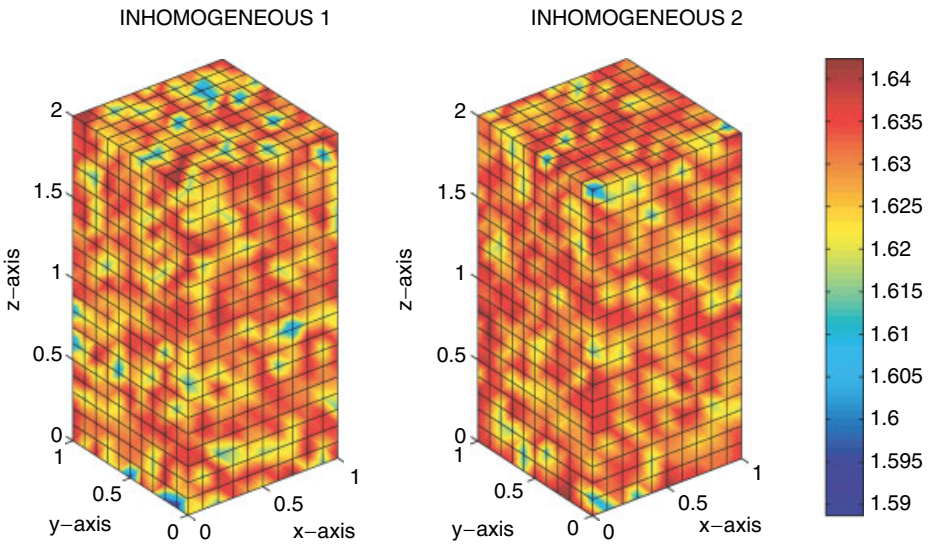

Figure 8. Initial specific volume field and finite element discretization for inhomogeneous rectangular specimens.

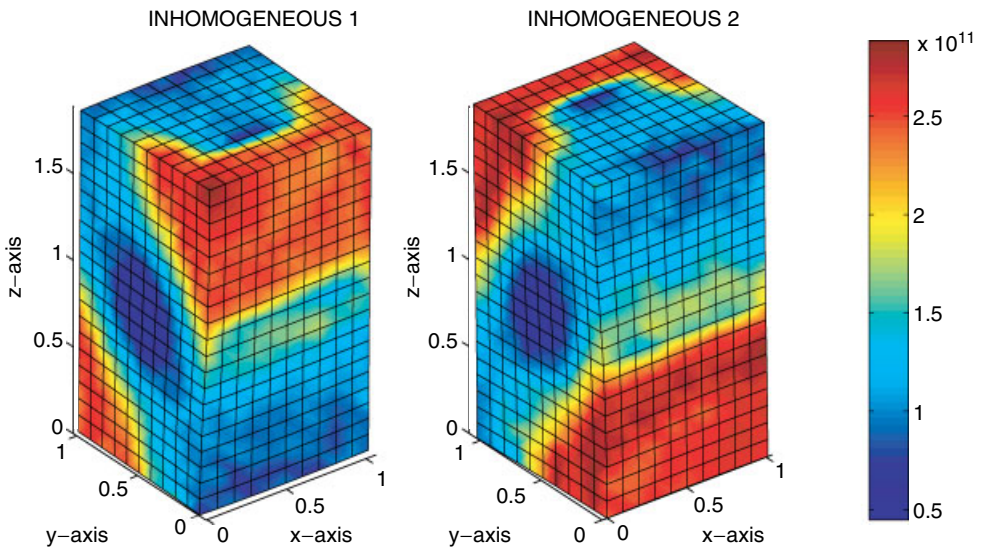

Figure 9. Contour of function $\mathscr{F}(\mathbf{A})$ at onset of localization for inhomogeneous rectangular specimens.

Figure 9 shows the contours of the function $\mathscr{F}(\mathbf{A})$, for both inhomogeneous samples, superimposed on the deformed mesh at the onset of localization. Here, we define localization when one or more Gauss points have detected the first negative incursion of the function

$\mathscr{F}(\mathbf{A})$ (even though contours appear as positive due to averaging necessary for plotting). Both inhomogeneous samples localized at a nominal axial strain of around 5.8%, yet, from Figure 9, we observe that the deformation patterns and the inclinations of the impending shear bands are conjugate to each other, almost looking as mirror images. This same trend is observed in Figure 10 which compares the contours of the total deviatoric strain invariant on both inhomogeneous samples. Comparing Figures 9 and 10, note a strong correlation between the vanishing of the determinant of the acoustic tensor and the localized values of the deviatoric component of deformation (shear strains).

Figure 11 shows the nominal axial stress as a function of the nominal axial strain for both inhomogeneous specimens and their homogeneous counterpart. We see a close agreement in the global responses up to about 4% nominal axial strain, where the inhomogeneous responses softened and eventually localized at 5.8% strain. The homogeneous sample, however, underwent very little softening and in fact continued to harden. The change in volume experienced by the heterogeneous and homogeneous samples are compared in Figure 12. Unlike Figure 11, the responses seem identical throughout the duration of the simulations. Note that the samples compacted first, and then dilated. This trend is very common in dense sands.

From these simulations we conclude that perturbations in the density field via inhomogeneities at the meso-scale tend to trigger the onset of localization. Also, as we compare the responses of two nearly identical inhomogeneous samples, we observe localization patterns that are conjugate of each another. The mechanical responses of the heterogeneous samples in the early stages of deformation are very similar to those of their homogeneous counterpart, suggesting that we can calibrate the meso-scale model parameters from global specimen responses during the early stages of loading.

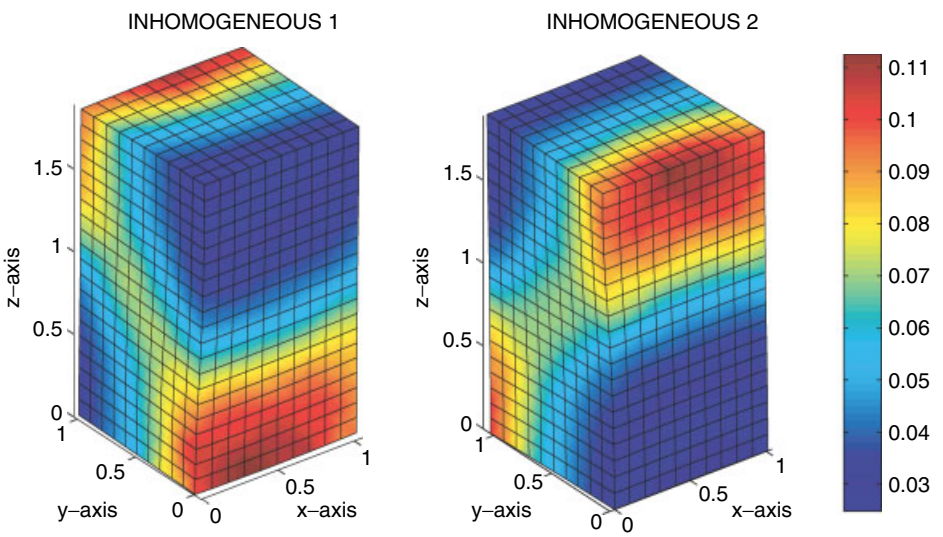

Figure 10. Deviatoric strain invariant field at onset of localization for inhomogeneous rectangular specimens.

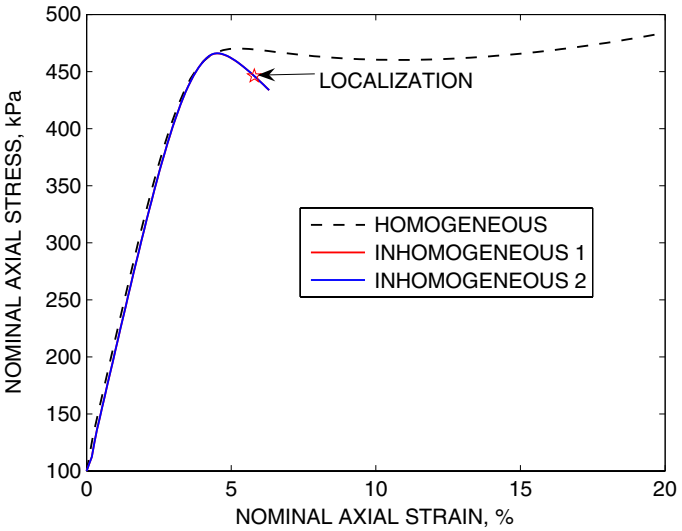

Figure 11. Nominal axial stress response for rectangular specimens.

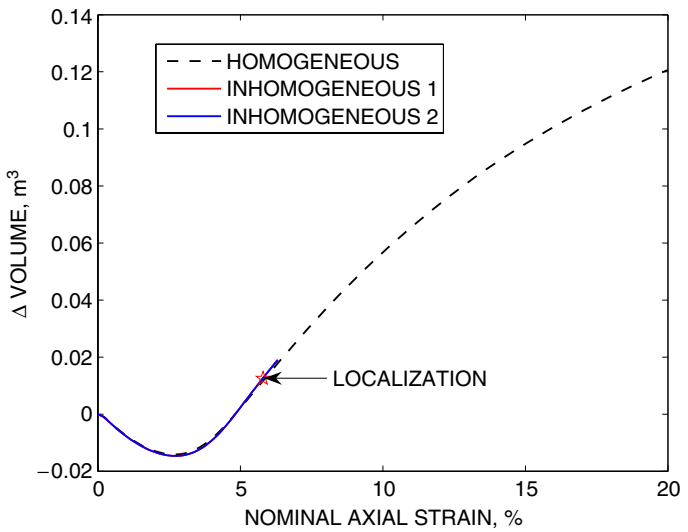

Figure 12. Volume change response for rectangular specimens.

6.3. Simulations with cylindrical specimens

Next, two inhomogeneous cylindrical specimens exhibiting random unstructured density fields were loaded in biaxial compression with BCs (and geometry) that emulated 'triaxial' testing in the laboratory. Both specimens are 2 m in diameter and 4 m tall. The elastoplastic properties

of these specimens are identical to those in the simulations with cubical specimens described above. However, here we used two different density fields to simulate the effects of increasing the spread in the random field on the localization properties of dense sand samples. To this end, the first inhomogeneous sample, which we call 'INHOMO 1.58-1.61', has a density field such that $\bar{e} = 0.59$, with lower and upper bounds $e_d = 0.58$ and $e_l = 0.61$, representing a relatively narrow distribution of void ratio. As for the second sample, we increased the spread in the void ratio field while keeping the mean constant, i.e. $\bar{e} = 0.59$, with lower and upper bounds $e_d = 0.56$ and $e_l = 0.68$. We will call this specimen 'INHOMO 1.56-1.68'. Figure 13 displays the initial specific volume field superimposed on the original configuration for both inhomogeneous samples. The cylindrical domains were discretized using 1280 eight-node brick elements equipped with the B-bar method.

As for the BCs, the specimens were biaxially loaded with a confining pressure applied on the sleeve of the cylinders (Neumann BC) while the top and bottom faces were supported on rollers (Dirichlet BC). The point $(0, 0, 0)$ m in the specimens was supported by a pin for stability. As in the rectangular specimen simulations, the bottom face was constrained to move vertically while the top face was compressed to mimic a homogeneous deformation. The samples were then loaded as before, with an initial confining pressure of 100 kPa followed by axial compression from the top face until localization was detected.

Figures 14 and 15 show an 'X-ray' rendering of the deviatoric strain invariant and the function $\mathscr{F}(\mathbf{A})$, respectively, superimposed on different cut-planes in the deformed configuration for the 'INHOMO 1.58-1.61' sample. The sub-figures on the right show longitudinal panels while the sub-figures on the left show axial planes. Note a localized area of intense shearing going from the top-right to the bottom-left corner, compared with the rest of the sample that experienced lower values of shear strains. Similar trends were observed on the 'INHOMO 1.56-1.68' sample.

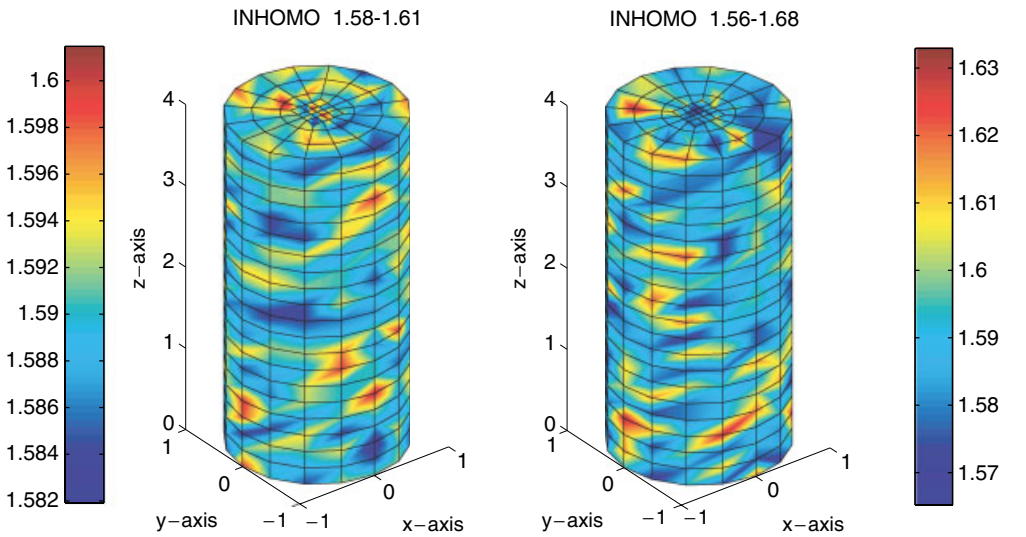

Figure 13. Initial specific volume field and finite element discretization for inhomogeneous cylindrical specimens.

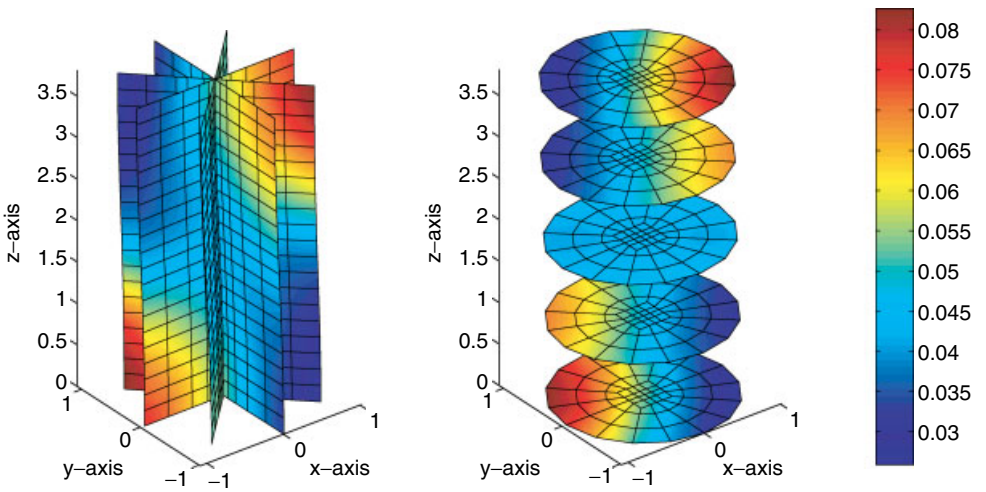

Figure 14. Total deviatoric strain invariant on various cut-planes at the onset of localization for sample 'INHOMO 1.58-1.61'.

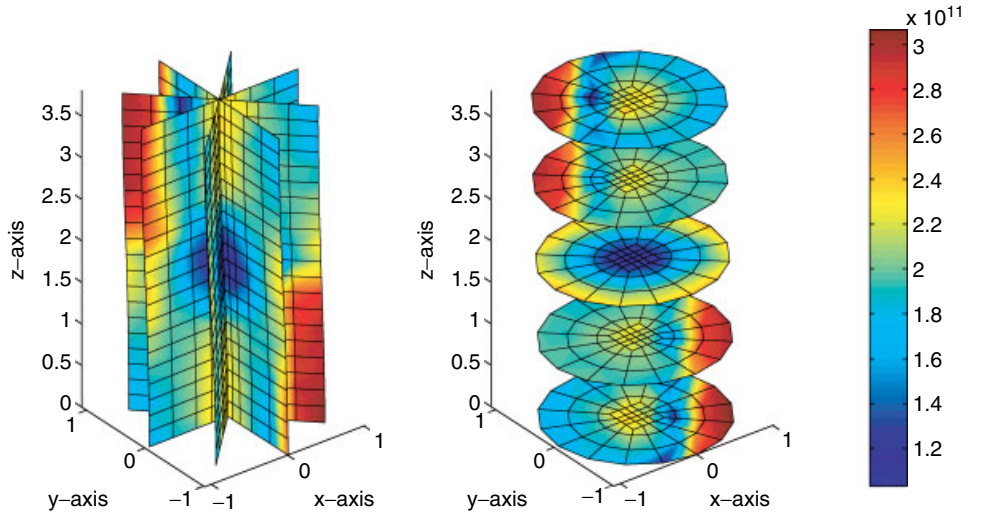

Figure 15. Determinant of acoustic tensor on various cut-planes at the onset of localization for sample 'INHOMO 1.58-1.61'.

Figures 16 and 17 show nearly identical plots of the overall axial stress–axial strain and volumetric strain–axial strain for both inhomogeneous samples and their homogeneous counterpart with $e = 1.59$. The inhomogeneous samples localized at different instants whereas the

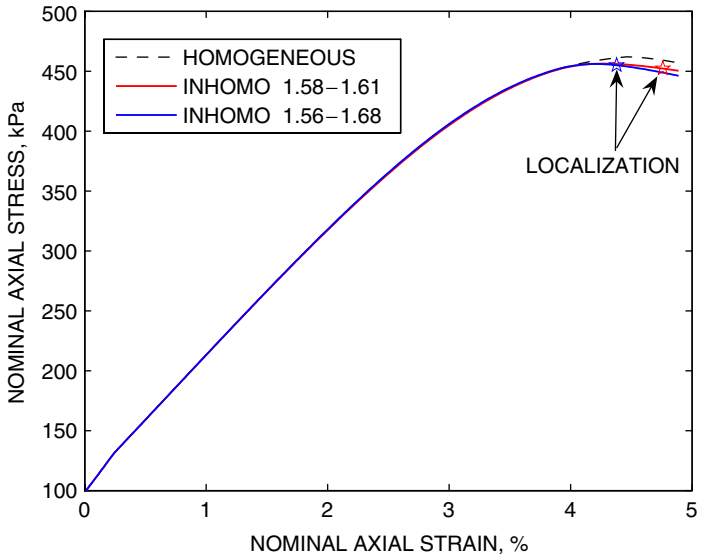

Figure 16. Nominal axial stress response for cylindrical specimens.

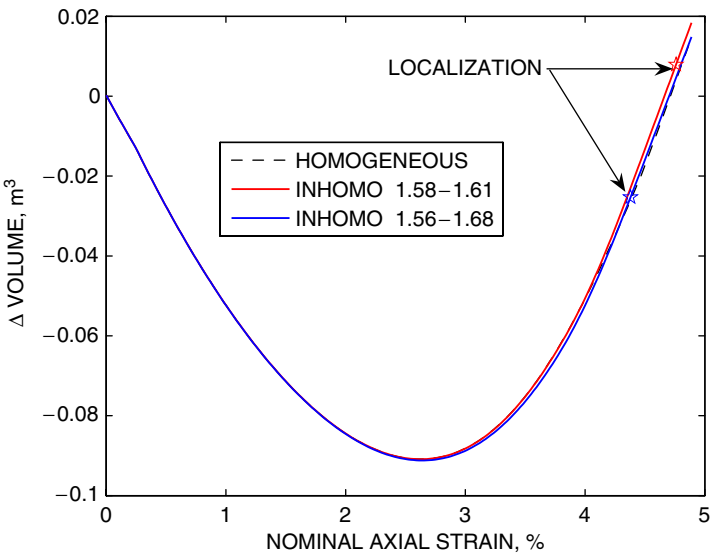

Figure 17. Volume change response for cylindrical specimens.

homogeneous sample did not localize at all. Sample 'INHOMO 1.56-1.68' localized at 4.38% whereas 'INHOMO 1.58-1.61' localized at 4.76%. The global convergence profile for the sample 'INHOMO 1.56-1.68' is reported in Figure 18, showing optimal rate of convergence.

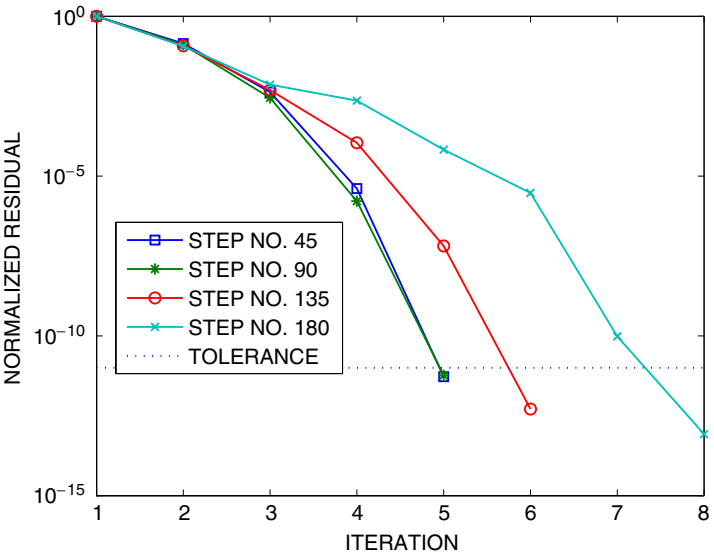

Figure 18. Convergence profile for finite element solution at various load steps.

7. CLOSURE

We have presented a critical state plasticity model that utilizes all three invariants of the stress tensor for deformation and strain localization analyses of granular materials with random unstructured density at the meso-scale. For the implementation, we have used the classical return mapping algorithm in the direction of the elastic logarithmic principal stretches. An iterative algorithm based on Newton's method was shown to deliver optimal performance. We also have presented a search algorithm for the minimum determinant of the acoustic tensor on the elastic principal stretch planes. Boundary-value problems mimicking soil samples have been analysed using a truncated exponential distribution to generate random and unstructured density fields. Results of numerical simulations suggest that inhomogeneities tend to trigger strain localization, with heterogeneous samples localizing when their equivalent homogeneous counterparts would not. Results of the studies also suggest that heterogeneity in the density field tends to enhance and accelerate the onset of strain localization, even if the initial global deformation responses of a homogeneous sample and its heterogeneous counterpart appear to be the same. The framework presented in this paper is useful for investigating strain localization phenomena in heterogeneous granular materials at a finer scale.

ACKNOWLEDGEMENTS

This work has been supported by National Science Foundation under Grant Nos. CMS-0201317 and CMS-0324674 to Stanford University. The first author is grateful to Professor Adrian Lew of Stanford University for enlightening discussions regarding constitutive theory and Lie derivatives. We also thank the two anonymous reviewers for their constructive reviews.

REFERENCES

1. Borja RI, Aydin A. Computational modeling of deformation bands in granular media. I. Geological and mathematical framework. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:2667–2698.
2. Desrues J, Viggiani G. Strain localization in sand: an overview of the experimental results obtained in grenoble using stereophotogrammetry. *International Journal for Numerical and Analytical Methods in Geomechanics* 2004; **28**:279–321.
3. Rechenmacher AL, Finno RJ. Shear band displacements and void ratio evolution to critical state in dilative sands. In *Bifurcation and Instabilities in Geomaterials*, Labuz JF, Drescher A (eds). A. A. Blakema Publishers: Rotterdam, 2003; 13–22.
4. Rechenmacher AL, Finno RJ. Digital image correlation to evaluate shear banding in dilative sands. *Geotechnical Testing Journal* 2004; **27**:13–22.
5. Jefferies MG. Nor-Sand: a simple critical state model for sand. *Géotechnique* 1993; **43**:91–103.
6. Borja RI, Andrade JE. Critical state plasticity, part VI: meso-scale finite element simulation of strain localization in discrete granular materials. *Computer Methods in Applied Mechanics and Engineering* 2006, in press (John Argyris Memorial special issue).
7. Been K, Jefferies MG. A state parameter for sands. *Géotechnique* 1985; **35**:99–112.
8. Manzari MT, Dafalias YF. A critical state two-surface plasticity model for sands. *Géotechnique* 1997; **43**:255–272.
9. Roscoe KH, Burland JH. On the generalized stress–strain behaviour of 'wet' clay. In *Engineering Plasticity*, Heyman J, Leckie FA (eds). Cambridge University Press: Cambridge, 1968; 535–609.
10. Borja RI, Lee SR. Cam–Clay plasticity, part I: implicit integration of elasto-plastic constitutive relations. *Computer Methods in Applied Mechanics and Engineering* 1990; **78**:49–72.
11. Borja RI. Cam–Clay plasticity, part II: implicit integration of constitutive equation based on nonlinear elastic stress predictor. *Computer Methods in Applied Mechanics and Engineering* 1991; **88**:225–240.
12. Borja RI, Tamagnini C. Cam–Clay plasticity, part III: extension of the infinitesimal model to include finite strains. *Computer Methods in Applied Mechanics and Engineering* 1998; **155**:73–95.
13. Peric D, Ayari MA. Influence of Lode's angle on the pore pressure generation in soils. *International Journal of Plasticity* 2002; **18**:1039–1059.
14. Peric D, Ayari MA. On the analytical solutions for the three-invariant Cam Clay model. *International Journal of Plasticity* 2002; **18**:1061–1082.
15. Borja RI, Sama KM, Sanz PF. On the numerical integration of three-invariant elastoplastic constitutive models. *Computer Methods in Applied Mechanics and Engineering* 2003; **192**:1227–1258.
16. Borja RI. Computational modeling of deformation bands in granular media. II. Numerical simulations. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:2699–2718.
17. Lode W. Versuche ueber den einfluss der mitt leren Hauptspannung auf das Fliessen der Metalle Eisen Kupfer und Nickel. *Zeitschrift fuer Physik* 1926; **36**:913–939.
18. Borja RI. Bifurcation of elastoplastic solids to shear band mode at finite strains. *Computer Methods in Applied Mechanics and Engineering* 2002; **191**:5287–5314.
19. Lee EH. Elastic-plastic deformation at finite strain. *Journal of Applied Mechanics* 1969; **36**:1–6.
20. Holzapfel GA. *Nonlinear Solid Mechanics*. Wiley: West Sussex, U.K., 2000.
21. Marsden JE, Hughes TJR. *Mathematical Theory of Elasticity*. Prentice-Hall: Englewood Cliffs, NJ, 1983.
22. Houlsby GT. The use of a variable shear modulus in elasto-plastic models for clays. *Computers and Geotechnics* 1985; **1**:3–13.
23. Borja RI, Tamagnini C, Amorosi A. Coupling plasticity and energy-conserving elasticity models for clays. *Journal of Geotechnical and Geoenvironmental Engineering* (ASCE) 1997; **123**:948–957.
24. Borja RI. *Plasticity Modeling and Computation*. Lecture Notes. Stanford University: California, 2005.
25. Gudehus G. Elastoplastische stoffgleichungen fur trockenen sand. *Ingenieur-Archiv* 1973; **42**:151–169.
26. Argyris JH, Faust G, Szimmat, J, Warnke EP, Willam KJ. Recent developments in the finite element analysis of prestressed concrete reactor vessels. *Nuclear Engineering and Design* 1974; **28**:42–75.
27. Jiang J, Pietruszczak. Convexity of yield loci for pressure sensitive materials. *Computers and Geotechnics* 1988; **5**:51–63.
28. Willam KJ, Warnke EP. Constitutive model for the triaxial behaviour of concrete. *Concrete Structures Subjected to Triaxial Stresses*, ISMES, Bergamo, Italy, May 1974.
29. Dafalias YF. Plastic spin: necessity or redundancy? *International Journal of Plasticity* 1998; **14**:909–931.
30. Simo JC, Hughes TJR. *Computational Inelasticity*. Prentice-Hall: New York, 1998.

31. Lade PV. Static instability and liquefaction of loose fine sandy slopes. *Journal of Geotechnical Engineering* (ASCE) 1992; **118**:51–71.
32. Lade PV, Duncan JM. Elastoplastic stress–strain theory for cohesionless soil. *Journal of the Geotechnical Engineering Division* (ASCE) 1975; **101**:1037–1053.
33. Simo JC. Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory. *Computer Methods in Applied Mechanics and Engineering* 1992; **99**:61–112.
34. Andrade JE, Borja RI. Fully implicit numerical integration of a hyperelastoplastic model for sands based on critical state plasticity. In *Computational Fluid and Solid Mechanics*, Bathe KJ (ed.). Elsevier Science Ltd.: Amsterdam, 2005; 52–54.
35. Bonet J, Wood RD. *Nonlinear Continuum Mechanics for Finite Element Analysis*. Cambridge University Press: Cambridge, U.K., 1997.
36. Belytschko T, Liu WK, Moran B. *Nonlinear Finite Elements for Continua and Structures*. Wiley: West Sussex, U.K., 2000.
37. Ortiz M, Martin JB. Symmetry-preserving return mapping algorithms and incrementally extremal paths—a unification of concepts. *International Journal for Numerical Methods in Engineering* 1989; **8**:1839–1853.
38. Rudnicki JW, Rice JR. Conditions for localization of deformation in pressure-sensitive dilatant materials. *Journal of the Mechanics and Physics of Solids* 1975; **23**:371–394.
39. Ortiz M, Leroy Y, Needleman A. A finite element method for localized failure analysis. *Computer Methods in Applied Mechanics and Engineering* 1987; **61**:189–214.
40. Willam K. Constitutive models for engineering materials. In *Encyclopedia of Physical Science and Technology*, vol. 3 (3rd edn). Academic Press: New York, 2002.
41. Oliver J, Huespe AE. Theoretical and computational issues in modeling material failure in strong discontinuity scenarios. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:2987–3014.
42. Ottosen NS, Runesson K. Properties of discontinuous bifurcation solutions in elasto-plasticity. *International Journal of Solids and Structures* 1991; **27**:401–421.
43. Mosler J. Numerical analyses of discontinuous material bifurcation: strong and weak discontinuities. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**:979–1000.
44. Shahinpoor M. Statistical mechanical considerations on the random packing of granular materials. *Powder Technology* 1980; **25**:163–176.
45. Nübel K, Huang W. A study of localized deformation pattern in granular media. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:2719–2743.
46. Benjamin JR, Cornell CA. *Probability*, *Statistics*, *and Decision for Civil Engineers*. McGraw-Hill, Inc.: New York, 1970.
47. Nagtegaal JC, Parks DM, Rice JR. On numerically accurate finite element solutions in the fully plastic regime. *Computer Methods in Applied Mechanics and Engineering* 1974; **4**:153–177.
48. Simo JC, Taylor RL, Pister KS. Variational and projection methods for the volume constraint in finite deformation elasto-plasticity. *Computer Methods in Applied Mechanics and Engineering* 1985; **51**:177–208.
49. Hughes TJR. *The Finite Element Method*. Prentice-Hall: Englewood Cliffs, NJ, 1987.

