



Critical state plasticity. Part VII: Triggering a shear band in variably saturated porous media



Ronaldo I. Borja^{a,*}, Xiaoyu Song^a, Wei Wu^b

^aDepartment of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA

^bInstitut für Geotechnik, Universität für Bodenkultur, Feistmantelstraße 4, 1180 Vienna, Austria

ARTICLE INFO

Article history:

Received 1 October 2012

Received in revised form 27 February 2013

Accepted 11 March 2013

Available online 4 April 2013

Keywords:

Bifurcation

Critical state theory

Heterogeneity

Mixed formulation

Shear band

Unsaturated porous media

ABSTRACT

In a previous paper (Part VI), the impact of spatially varying density on the localization of deformation of granular materials has been quantified using mesoscopic representations of stresses and deformation. In the present paper, we extend the formulation to unsaturated porous media and investigate the effect of spatially varying degree of saturation on triggering a shear band in granular materials. Variational formulations are presented for porous solids whose voids are filled with liquid and gas. Two critical state formulations are used to characterize the solid constitutive response: one developed for clay and another for sand. Stabilized low-order mixed finite elements are used to solve the fully coupled solid-deformation/fluid-flow problem. For the first time, we present the consistent derivative of the effective stress tensor with respect to capillary pressure considering full coupling of solid deformation with fluid flow, which is essential for achieving an optimal convergence rate of Newton iteration.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

A class of problems that has attracted considerable attention in computational solid mechanics involves very large deformation occurring over a very narrow zone. Deformation bands are narrow zones of intense shear, compaction, and/or dilation; the displacement field is continuous but the strain field exhibits a discontinuity [15,16]. Material and/or geometric imperfection is known to be a common trigger of deformation bands [8,9]. In the past, arbitrary imperfections in the form of weak elements have been used in finite element simulations to trigger strain localization because of the uncertainties in quantifying actual specimen imperfections. However, advances in nondestructive, noninvasive imaging techniques have now allowed for more accurate quantification of the spatial variation of density in a specimen of granular materials. It is well known that the strength and stiffness of a granular material correlate very well with density, so knowing the spatial variation of density allows the spatial inhomogeneities within a specimen to be prescribed deterministically. Density within a specimen is a continuum variable associated with the so-called ‘mesoscopic’ scale, a scale larger than the grains but smaller than the specimen (see [7] for a detailed description of the macro, meso, and grain scales). A mesoscopic characterization of inhomogeneity in a

specimen, be it in the form of density or some other continuum variables, is the point of view taken in this paper.

Apart from density, the degree of saturation representing the amount of water present in the pores of a material is another mesoscopic continuum variable that can potentially serve as an imperfection triggering strain localization. Conventionally denoted by the symbol S_r , the degree of saturation is known to influence the strength and permeability of a porous material such as soil. Typically, the degree of saturation is determined in the laboratory by taking the weight of a sample before and after drying, but this technique is destructive and can only describe an average value for the entire specimen but not the spatial mesoscopic distribution within the specimen. More recently, it has been shown that, like density, the degree of saturation can also be quantified nondestructively through imaging techniques along with digital image processing to allow deterministic characterization of its distribution within the specimen. Such finer-scale measurements of degree of saturation are critical for the mesoscale modeling technique advocated in this paper.

As a brief literature review, a variety of nondestructive, noninvasive laboratory techniques are currently utilized for the measurement of liquid saturation in porous media. They include gamma ray or conventional X-ray attenuation techniques [71]. In principle, these techniques exploit differences in the absorbance of electromagnetic energy between the liquid, gas and solid phases. Recently, synchrotron X-ray measurements have been developed as a reliable method for measuring phase saturation during multi-

* Corresponding author. Tel.: 1 650 7233664; fax: 1 650 7237514.

E-mail address: borja@stanford.edu (R.I. Borja).

phase transient flow [33,39,62,72]. The technique allows measurements with short counting time, but only regions less than 0.5 cm^2 can be characterized at a given time [32]. Image analysis methods have been valuable alternative tools in measuring transient phenomena in the entire flow domain. They have been used in miscible and immiscible experiments where various parameters linked to reflected light intensity recorded onto color or black and white photographs, and subsequently digitally scanned to be computer-processed, have been correlated to species concentration or liquid saturation [1,44,65,66,74].

Very recently, Yoshimoto et al. [78] proposed a method to directly measure the degree of saturation on a region by noting the variation in color of the ground with changes in the moisture content of the soil. They showed that the relation between degree of saturation and luminance value can be expressed in terms of a quadratic correlation function. With this method, contours of degree of saturation can be generated, making it possible to visualize the propagation of the saturated region (see Fig. 1). The idea is similar to the technique proposed by Darnault et al. [32], which is a variation of the method by Glass et al. [43] for air–water systems that uses light transmission method (LTM) to allow full field moisture content visualization in soil–oil–water systems. By appropriately coloring the water, they found the hue of the transmitted light to be directly related to the water content within the porous medium. To obtain the calibration curve between the hue value and oil–water content, they constructed a two-dimensional calibration chamber consisting of compartments with known quantities of oil and water, from which they concluded that a unique relationship exists between the hue and water content.

Kechavarzi et al. [49] developed a multispectral image analysis technique to determine dynamic distributions of non-aqueous phase liquids (NAPL), water, and air saturations in two-dimensional three-fluid phase laboratory experiments. They showed that the optical density for the reflected luminous intensity is a linear function of the NAPL and the water saturation for each spectral band and for any two- and three-fluid phase systems. This method was used to obtain a continuous, quantitative and dynamic full field mapping of the NAPL saturation as well as the variation of the water and the air saturation during NAPL flow. To summarize, a variety of nondestructive, noninvasive techniques for quantifying the spatial variation of degree of saturation are currently available,

in addition to similar techniques for quantifying the spatial variation of density (see [7,31,60] for a survey of the latter techniques).

This paper focuses on the degree of saturation as a trigger to strain localization in granular materials. As the water content of a porous medium increases, the apparent preconsolidation pressure decreases [38,54,58]. This implies that increasing the water content unevenly could induce nonuniform yielding in the material even without a change in the external load. We expect regions with a higher degree of saturation as likely hotspots for early yielding and for early onset of localized deformation. However, the degree of saturation is not the only possible source of material imperfection. As noted earlier, a spatially varying density could also be an important trigger of strain localization. In fact, it has been observed from previous numerical simulations and experiments that regions of high porosity are also likely hotspots for localized deformation in granular materials [3,7,11]. Given that degree of saturation and density are two independent state variables, they serve as independent sources of imperfection triggering strain localization in granular materials.

Density describes the state of the solid phase, whereas degree of saturation is a fluid state variable. Interaction between these two sources of material imperfection requires a fully coupled hydromechanical formulation [8,9,26,27,36,47,56,73,76,79]. We present a variational formulation for fully coupled solid deformation–fluid flow in unsaturated porous media for deformation and strain localization analyses. Important contributions of this paper include casting a nonstandard critical state model for sand [2,11,48] within the framework of the hydromechanical continuum theory, and an implicit implementation of the variational equations in the framework of mixed finite element formulation. We remark that conventional critical state plasticity models, including the modified Cam-Clay model [14,20–22,63], cannot represent density imperfection since these models uniquely determine the void ratio from the critical state line and the current state of stress. A nonconventional critical state model that uses density as a principal state variable is the ‘Nor-Sand’ model [48]. We use this model to ‘uncouple’ the void ratio from the critical state line, allowing a spatially varying density to be specified independent of the state of stress. We show how the computational framework presented in this paper accommodates spatially varying density and degree of saturation simultaneously.

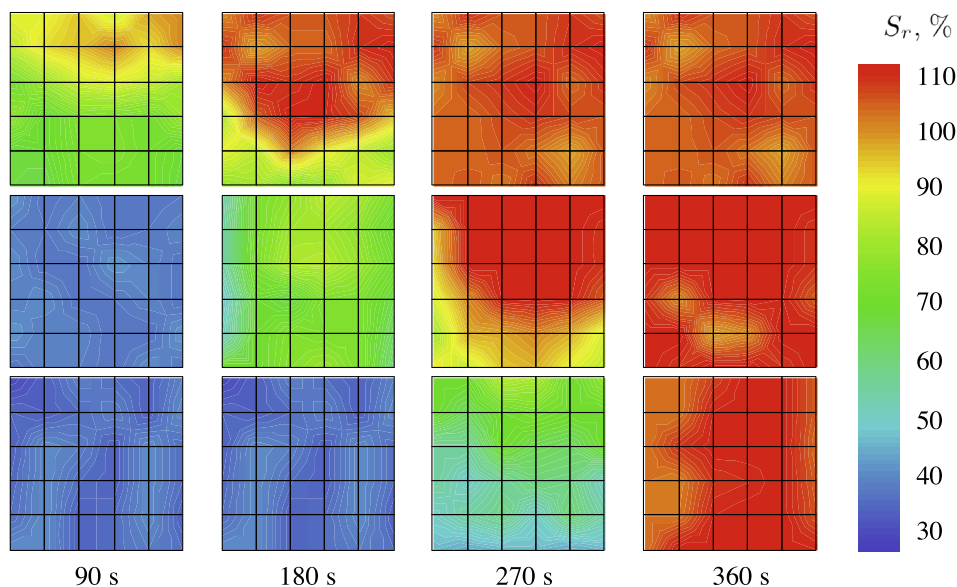


Fig. 1. Contour of degree of saturation as a function of time for Toyoura sand with imposed fluid head of 2 cm on top and drained at the bottom. Each box with 25 cells is $10 \text{ cm} \times 10 \text{ cm}$. Color bar indicates an error within 10% as indicated by the value $S_r = 110\%$, which is 10% in excess of the theoretical maximum value. After Reference [78].

The scope of the paper is limited to the triggering of a persistent shear band in variably saturated porous media. Once a persistent shear band has been identified, post-localization enhancements, either through the assumed enhanced strain or extended finite element methods, can be employed to capture the evolution of the identified shear band at post-failure condition [18,19,26,51–53,61]. We also limit the paper to a deterministic representation of the spatial variability of density and degree of saturation, which we assume can be measured and quantified in the laboratory. The formulation advanced in the paper can be used to provide a mechanistic underpinning for any uncertainty or probabilistic model, although such application is beyond the scope of the paper (see [2,29,68,70] for a sampling of stochastic simulations in geomechanics).

2. Variational formulation

We consider a mixture of solid matrix with continuous voids filled with water and air. The total volume of the mixture is $V = V_s + V_w + V_a$ and the total mass is $M = M_s + M_w + M_a$, where $M_\alpha = \rho_\alpha V_\alpha$ for $\alpha = \text{solid, water, and air}$; and ρ_α is the true mass density of the α constituent. The volume fraction occupied by the α constituent is given by $\phi_\alpha = V_\alpha/V$, so that

$$\phi^s + \phi^w + \phi^a = 1. \quad (1)$$

The partial mass density of the α constituent is given by $\rho^\alpha = \phi^\alpha \rho_\alpha$, where ρ_α is the intrinsic mass density of the α constituent. This gives

$$\rho^s + \rho^w + \rho^a = \rho, \quad (2)$$

where $\rho = M/V$ is the total mass density of the mixture.

We define void fractions S_r and $1 - S_r$ representing the portions of void occupied by water and air, respectively. The void fractions are related to the volume fractions through the equations

$$S_r = \frac{\phi^w}{1 - \phi^s}, \quad 1 - S_r = \frac{\phi^a}{1 - \phi^s}. \quad (3)$$

The void fraction S_r is called the degree of saturation and is used in the effective stress equation [10,12]

$$\boldsymbol{\sigma} = \bar{\boldsymbol{\sigma}} - Bp^* \mathbf{1}, \quad p^* = S_r p + (1 - S_r)p_a, \quad (4)$$

where $\boldsymbol{\sigma}$ and $\bar{\boldsymbol{\sigma}}$ are the total and effective Cauchy stress tensors, respectively, p and p_a are the pore water and pore air pressures, $\mathbf{1}$ is the second-order identity tensor, and B is the Biot coefficient. For soils, $B = 1$ is a reasonable approximation. In this paper, we assume that $p_a = 0$ (atmospheric pressure), and that the process is isothermal (see [35,55], for example, on how to include thermal effects).

We consider a partially saturated mixture in domain \mathcal{B} with boundary $\partial\mathcal{B} = \overline{\partial\mathcal{B}_u} \cup \overline{\partial\mathcal{B}_t}$, where $\partial\mathcal{B}_u$ and $\partial\mathcal{B}_t$ are nonintersecting portions of the total boundary $\partial\mathcal{B}$ on which the solid displacements and total tractions, respectively, are prescribed. Ignoring inertia loads (see Uzuoka and Borja [73] for a formulation with inertia load), the momentum conservation equation along with relevant boundary conditions can be stated as follows. Find \mathbf{u} and p such that

$$\nabla \cdot (\bar{\boldsymbol{\sigma}} - S_r p \mathbf{1}) + \rho \mathbf{g} = \mathbf{0} \quad \text{in } \mathcal{B}, \quad (5)$$

subject to boundary conditions

$$\mathbf{u} = \hat{\mathbf{u}} \quad \text{on } \partial\mathcal{B}_u \quad \text{and} \quad \mathbf{n} \cdot \boldsymbol{\sigma} = \hat{\mathbf{t}} \quad \text{on } \partial\mathcal{B}_t, \quad (6)$$

where \mathbf{g} is the gravity acceleration vector, \mathbf{n} is the outward unit normal vector to the boundary, and $\hat{\mathbf{u}}$ and $\hat{\mathbf{t}}$ are given space and time functions.

We next decompose the same boundary into $\partial\mathcal{B} = \overline{\partial\mathcal{B}_p} \cup \overline{\partial\mathcal{B}_q}$, where $\partial\mathcal{B}_p$ and $\partial\mathcal{B}_q$ are nonintersecting portions of the total boundary $\partial\mathcal{B}$ on which the pore water pressure and fluid flux, respectively, are prescribed. The mass conservation equations for water along with relevant fluid flow boundary conditions can be stated as follows. Find \mathbf{u} and p such that

$$(1 - \phi^s) \dot{S}_r + \frac{\phi^w}{K_w} \dot{p} + S_r \nabla \cdot \mathbf{v} = -\frac{1}{\rho_w} \nabla \cdot \mathbf{w}, \quad (7)$$

subject to pressure \hat{p} and flux \hat{q} boundary conditions

$$p = \hat{p} \quad \text{on } \partial\mathcal{B}_p \quad \text{and} \quad \mathbf{n} \cdot (\phi^w \tilde{\mathbf{v}}) = -\hat{q} \quad \text{on } \partial\mathcal{B}_q, \quad (8)$$

where $\mathbf{v} = \dot{\mathbf{u}}$ is the velocity of the solid; \mathbf{w} is the Eulerian relative water flow vector given by

$$\mathbf{w} = \rho_w \phi^w \tilde{\mathbf{v}}, \quad \tilde{\mathbf{v}} = \mathbf{v}_w - \mathbf{v} \quad (9)$$

is the relative velocity of water with respect to the solid, K_w is the bulk modulus of water, and the superimposed dot denotes a material time derivative following the motion of the solid. We assume that water is incompressible and ignore the second term on the left-hand side of (7) for simplicity, and set $\rho_w = \text{constant}$. The product term $\phi^w \tilde{\mathbf{v}} \equiv \bar{\mathbf{v}}$ is known as Darcy velocity.

Relevant constitutive laws motivate the \mathbf{u}/p formulation implied above. Take, for example, the degree of saturation S_r that is related to the suction stress $-p$ through the water retention curve, Darcy velocity $\bar{\mathbf{v}}$ that is related to the pressure gradient ∇p via Darcy's law, and the effective Cauchy stress tensor $\bar{\boldsymbol{\sigma}}$ that is related to the infinitesimal strain rate tensor $\nabla^s \mathbf{v} \equiv (\nabla \mathbf{v} + \mathbf{v} \nabla)/2$ and (indirectly) to the suction stress $-p$ through an elastoplastic constitutive law. The independent variables then boil down to \mathbf{u} and p . In the following section we consider two critical-state constitutive laws, one typically associated with clays and the other with sands, and cast both of them within the framework of the \mathbf{u}/p formulation.

The variational equation for linear momentum balance takes the form

$$\int_{\mathcal{B}} \nabla^s \boldsymbol{\omega} : (\bar{\boldsymbol{\sigma}} - S_r p \mathbf{1}) dV = \int_{\mathcal{B}} \boldsymbol{\omega} \cdot \rho \mathbf{g} dV + \int_{\partial\mathcal{B}_t} \boldsymbol{\omega} \cdot \hat{\mathbf{t}} dA, \quad (10)$$

where $\boldsymbol{\omega}$ is the vector of displacement variation such that $\omega_i \in H^1$ and $\omega_i = 0$ on $\partial\mathcal{B}_{u_i}$, and ∇^s denotes the symmetric component of the gradient operator. Similarly, the variational equation for fluid flow can be written as

$$\int_{\mathcal{B}} \theta S_r \nabla \cdot \mathbf{v} dV + \int_{\mathcal{B}} \theta (1 - \phi^s) \dot{S}_r dV - \int_{\mathcal{B}} \nabla \theta \cdot \bar{\mathbf{v}} dV = \int_{\partial\mathcal{B}_q} \theta \hat{q} dA, \quad (11)$$

where θ is the pressure variation such that $\theta \in H^1$ and $\theta = 0$ on $\partial\mathcal{B}_p$. We can integrate them in time to obtain the discrete evolutions of \mathbf{u} and p . In so doing, we assume that \mathbf{u}_n and p_n are known at time t_n and we want to determine \mathbf{u} and p at time $t = t_n + \Delta t$ (the usual subscript ' $n + 1$ ' is dropped for brevity). Eq. (10) is an elliptic equation for which time integration is straightforward,

$$\mathcal{L} = \int_{\mathcal{B}} \nabla^s \boldsymbol{\omega} : (\bar{\boldsymbol{\sigma}} - S_r p \mathbf{1}) dV - \int_{\mathcal{B}} \boldsymbol{\omega} \cdot \rho \mathbf{g} dV - \int_{\partial\mathcal{B}_t} \boldsymbol{\omega} \cdot \hat{\mathbf{t}} dA, \quad (12)$$

where all of the variables are assumed to be evaluated at time t . The conservation of momentum is then given simply by the condition $\mathcal{L} = 0$. However, Eq. (11) is a nonlinear first-order equation for which the generalized trapezoidal time-integration method would prove challenging to implement.

Consider the time integration of (11) by the one-parameter generalized trapezoidal method such that the time integration parameter is $\beta = 1$ for backward implicit and $\beta = 0$ for forward Euler. The integrated variational equation takes the form

$$\begin{aligned} \widetilde{\mathcal{M}} = & \Delta t \int_B \theta f_{n+\beta} dV + \int_B \theta (1 - \phi^s) (S_r - S_{r,n}) dV \\ & - \Delta t \int_B \nabla \theta \cdot \bar{\mathbf{v}}_{n+\beta} dV - \Delta t \int_{\partial B_q} \theta \widehat{q}_{n+\beta} dA, \end{aligned} \quad (13)$$

where

$$f_{n+\beta} = \beta S_r \nabla \cdot \dot{\mathbf{u}} + (1 - \beta) S_{r,n} \nabla \cdot \dot{\mathbf{u}}_n \quad (14)$$

$$\bar{\mathbf{v}}_{n+\beta} = \beta \bar{\mathbf{v}} + (1 - \beta) \bar{\mathbf{v}}_n \quad (14)$$

$$\widehat{q}_{n+\beta} = \beta \widehat{q} + (1 - \beta) \widehat{q}_n. \quad (15)$$

We assume that the change in ϕ^s is small, since in variably saturated porous media the contribution of the change in porosity to storage properties is not as significant as the change in degree of saturation [25]. Given that S_r and $\bar{\mathbf{v}}$ vary nonlinearly with p , any method other than $\beta = 1$ would be difficult to implement, see [28].

Motivated by the return mapping algorithm that employs the standard backward implicit scheme for stress-point integration, we take $\beta = 1$ and write the conservation of water mass as

$$\begin{aligned} \mathcal{M} = & \int_B \theta S_r \nabla \cdot (\mathbf{u} - \mathbf{u}_n) dV + \int_B \theta (1 - \phi^s) (S_r - S_{r,n}) dV \\ & - \Delta t \int_B \nabla \theta \cdot \bar{\mathbf{v}} dV - \Delta t \int_{\partial B_q} \theta \widehat{q} dA. \end{aligned} \quad (16)$$

Hence, the problem is to find \mathbf{u} and p such that $\mathcal{L} = \mathcal{M} = 0$. Given that both the fluid flow and solid deformation are integrated consistently by the backward implicit scheme, it seems plausible to expect that the accuracy and (linearized) stability properties of the method are preserved, see [40].

Next, we quantify the variation of \mathcal{L} and \mathcal{M} to variations in \mathbf{u} and p . Apart from the fact that this would reveal the intricate coupling between the two independent variables in the unsaturated regime, the result is also useful for constructing the algorithmic tangent operator in Newton iteration. The variation of \mathcal{L} is given by

$$\delta \mathcal{L} = \int_B \nabla^s \boldsymbol{\omega} : (\mathbf{c} : \nabla^s \delta \mathbf{u} + \mathbf{a} \delta p - \bar{S}_r \delta p \mathbf{1}) dV, \quad (17)$$

where

$$\mathbf{c} = \frac{\partial \bar{\boldsymbol{\sigma}}}{\partial \boldsymbol{\epsilon}}, \quad \mathbf{a} = \frac{\partial \bar{\boldsymbol{\sigma}}}{\partial p}, \quad \bar{S}_r = S_r + p S'_r(p) \quad (18)$$

and $\boldsymbol{\epsilon} = \nabla^s \mathbf{u}$ is the infinitesimal strain tensor. We refer to [14] for further details of the above equations. In the above definitions, \mathbf{c} is the usual algorithmic stress-strain tensor and S_r is related to suction stress $-p$ through the water retention curve. The tensor \mathbf{a} is unique to the unsaturated porous media formulation in that it reflects dependence of the calculated effective stress $\bar{\boldsymbol{\sigma}}$ on the suction stress $-p$ through the so-called preconsolidation stress of the material.

Next, consider the following variation of \mathcal{M} for a fixed surface flux:

$$\begin{aligned} \delta \mathcal{M} = & \int_B \theta S_r \nabla \cdot \delta \mathbf{u} dV + \int_B \theta \nabla \cdot (\mathbf{u} - \mathbf{u}_n) S'_r(p) \delta p dV \\ & + \int_B \theta (1 - \phi^s) S'_r(p) \delta p dV - \Delta t \int_B \nabla \theta \cdot \delta \bar{\mathbf{v}} dV, \end{aligned} \quad (19)$$

where $\delta \bar{\mathbf{v}}$ is the variation of Darcy velocity. This variation can be obtained with the aid of Darcy's law, which takes the form

$$\bar{\mathbf{v}} = -k_{rw}(p) \mathbf{K}_{sat} \cdot \nabla \left(\frac{p}{\rho_w g} + z \right), \quad (20)$$

where \mathbf{K}_{sat} is the saturated hydraulic conductivity, g is the gravity acceleration constant, z is the vertical coordinate, and $k_{rw}(p)$ is the relative permeability that depends on the capillary pressure $-p$ in the unsaturated regime. Taking the variation gives

$$\delta \bar{\mathbf{v}} = - \frac{k_{rw}(p)}{\rho_w g} \mathbf{K}_{sat} \cdot \nabla (\delta p) - k'_{rw}(p) \delta p \mathbf{K}_{sat} \cdot \nabla \left(\frac{p}{\rho_w g} + z \right). \quad (21)$$

A specific soil-water characteristic curve that facilitates the evaluation of $k'_{rw}(p)$ is given in the next section.

3. Finite element formulation

We employ a mixed finite element formulation with equal order interpolation for displacement and pressures. The independent variables are the nodal displacements d_{Ai} and pressure p_A . Each node has $n_{sd} + 1$ degrees of freedom, where n_{sd} = number of spatial dimensions. Let N_A = global shape function for node A ; then

$$u_i(\mathbf{x}, t) = \sum_A N_A(\mathbf{x}) d_{Ai}(t), \quad p(\mathbf{x}, t) = \sum_A N_A(\mathbf{x}) p_A(t), \quad (22)$$

where $d_{Ai}(t)$ and $p_A(t)$ are the time-varying displacement component i and pressure at any node A , including those where the essential boundary conditions are specified. This makes the degree of saturation a dependent variable that is calculated from the pressure p through the water retention curve, which is expressed in functional form as as

$$S_r(\mathbf{x}, t) = S_r(p(\mathbf{x}, t)) = S_r \left(\sum_A N_A(\mathbf{x}) p_A(t) \right). \quad (23)$$

The symmetric part of the displacement gradient defines the infinitesimal strain tensor,

$$\epsilon_{ij} = \frac{1}{2} \left(\sum_A N_{A,j}(\mathbf{x}) d_{Ai}(t) + \sum_A N_{A,i}(\mathbf{x}) d_{Aj}(t) \right), \quad (24)$$

with trace

$$\epsilon_{ii} = \sum_A N_{A,i}(\mathbf{x}) d_{Ai}(t). \quad (25)$$

The gradient of the pressure field takes a similar form,

$$p_{,i}(\mathbf{x}, t) = \sum_A N_{A,i}(\mathbf{x}) p_A(t). \quad (26)$$

Two shape function matrices can be constructed from the same shape functions N_A . The first pertains to the scalar pressure field,

$$\bar{\mathbf{N}} = [N_1 \quad N_2 \quad \dots \quad N_n], \quad (27)$$

with corresponding gradient

$$\mathbf{E} = \nabla \bar{\mathbf{N}} = \begin{bmatrix} N_{1,x} & N_{2,x} & \dots & N_{n,x} \\ N_{1,y} & N_{2,y} & \dots & N_{n,y} \end{bmatrix}. \quad (28)$$

The second matrix pertains to the displacement vector field,

$$\mathbf{N} = \begin{bmatrix} N_1 & 0 & \dots & N_n & 0 \\ 0 & N_1 & \dots & 0 & N_n \end{bmatrix}, \quad (29)$$

with corresponding symmetric component of the gradient

$$\mathbf{B} = \nabla^s \mathbf{N} = [\mathbf{B}_1 \quad \mathbf{B}_2 \quad \dots \quad \mathbf{B}_n], \quad \mathbf{B}_a = \begin{bmatrix} N_{a,x} & 0 \\ 0 & N_{a,y} \\ N_{a,y} & N_{a,x} \end{bmatrix} \quad (30)$$

The finite element momentum balance equation then takes the vector form

$$\int_B \mathbf{B}^T \{ \bar{\boldsymbol{\sigma}} - S_r p \mathbf{1} \} dV = \int_B \mathbf{N}^T \rho \mathbf{g} dV + \int_{\partial B_t} \mathbf{N}^T \hat{\mathbf{t}} dA, \quad (31)$$

whereas the time-integrated fluid mass balance equation is given by

$$\begin{aligned} & \int_B \bar{\mathbf{N}}^T S_r \nabla \cdot (\mathbf{u} - \mathbf{u}_n) dV + \int_B \bar{\mathbf{N}}^T (1 - \phi^s) (S_r - S_{r,n}) dV - \Delta t \int_B \mathbf{E}^T \bar{\mathbf{v}} dV \\ & = \Delta t \int_{\partial B_q} \bar{\mathbf{N}}^T \hat{\mathbf{q}} dA. \end{aligned} \quad (32)$$

Simultaneous solution of the above equations for the unknown nodal displacement vector \mathbf{d} and pressure vector \mathbf{p} necessitates an iterative strategy and consistent tangent operators. The next section elaborates the relevant tangent operators.

In the fully saturated regime $S_r = 1.0$, and the time-integrated fluid mass balance equation simplifies to the form

$$\int_B \bar{\mathbf{N}}^T \nabla \cdot (\mathbf{u} - \mathbf{u}_n) dV - \Delta t \int_B \mathbf{E}^T \bar{\mathbf{v}} dV = \Delta t \int_{\partial B_q} \bar{\mathbf{N}}^T \hat{\mathbf{q}} dA. \quad (33)$$

A very small Δt captures the incompressibility condition, which is given by

$$\int_B \bar{\mathbf{N}}^T \nabla \cdot (\mathbf{u} - \mathbf{u}_n) dV = 0. \quad (34)$$

In the incompressible and nearly incompressible regimes, equal-order interpolation for the displacement and pressure fields is known to cause spurious oscillation in the pressure field, unless some form of stabilization is utilized. Here, we employ the polynomial pressure projection stabilization advocated in [5,6,23,34] for Darcy and Stokes equations, and in [75] for the coupled solid deformation–fluid flow problem. This stabilization “corrects” the quantified deficiency of the linear-order pair, and is sufficiently robust for the numerical problems discussed in this paper.

4. Tangent operators

In this section, we derive general expressions for the tangent operators for any standard elastoplastic constitutive model in which the yield stress is also a function of the suction stress $s = -p$, which is the case for unsaturated porous materials. Two specific critical state models are then presented, one appropriate for clay and one for sand. Given that the elastic component of the constitutive model may also introduce nonlinear effects, the expressions are formulated in the elastic strain space. The development presented below accommodates any hyperelastic–plastic theory including those where the elastic bulk and shear moduli depend on the stress.

4.1. General expressions for solid tangent operators

We consider the standard return mapping algorithm in computational plasticity where $\bar{\boldsymbol{\sigma}}_n$ and $\Delta \epsilon$ are given and the stress tensor $\bar{\boldsymbol{\sigma}}$ at time $t = t_n + \Delta t$ is computed. For unsaturated porous media an additional variable that functions like a strain tensor, the incremental suction stress Δs , is also given, potentially affecting the final value of the effective stress $\bar{\boldsymbol{\sigma}}$. We define predictor (trial) values $\boldsymbol{\epsilon}^{e \text{ tr}} = \boldsymbol{\epsilon}_n^e + \Delta \boldsymbol{\epsilon}$, $s^{\text{tr}} = s_n + \Delta s$.

Note that for elastic process $\Delta \epsilon = \Delta \boldsymbol{\epsilon}^e$.

We define

$$\mathbf{x} = \begin{Bmatrix} \boldsymbol{\epsilon}^e \\ \Delta \lambda \end{Bmatrix}, \quad \mathbf{z} = \begin{Bmatrix} \boldsymbol{\epsilon}^{e \text{ tr}} \\ s^{\text{tr}} \end{Bmatrix}, \quad (36)$$

where $\Delta \lambda \geq 0$ is the standard incremental plastic multiplier. We can think of \mathbf{x} as a vector in \mathbb{R}^7 consisting of six components of the elastic strain tensor and a plastic multiplier, and \mathbf{z} as a vector consisting of the predictor elastic strain components and the suction stress. It is important to understand the following setup of the formulation: \mathbf{x} contains the local independent variables that satisfy the constitutive laws for a given \mathbf{z} , whereas \mathbf{z} contains the global

independent variables that satisfy the relevant conservation laws. In other words, $\mathbf{x} = \mathbf{x}(\mathbf{z})$.

In a local stress–point integration algorithm, \mathbf{z} is given and the task of the algorithm is to determine \mathbf{x} . Once $\boldsymbol{\epsilon}^e$ and $\Delta \lambda$ have been determined on the local level, the remaining state variables may be calculated from the following relations

$$\begin{aligned} \bar{\boldsymbol{\sigma}} &= \bar{\boldsymbol{\sigma}}(\boldsymbol{\epsilon}^e), \quad \bar{p}_c = \bar{p}_c(\bar{\boldsymbol{\sigma}}, \Delta \lambda, \boldsymbol{\epsilon}^{e \text{ tr}}, s^{\text{tr}}), \\ F &= F(\bar{\boldsymbol{\sigma}}, \bar{p}_c), \quad Q = Q(\bar{\boldsymbol{\sigma}}, \bar{p}_c), \end{aligned} \quad (37)$$

where F and Q are the yield and plastic potential functions, respectively, and $\bar{p}_c < 0$ is the preconsolidation stress.

A return mapping algorithm in elastic strain space may be employed to determine \mathbf{x} by defining the residual vector

$$\mathbf{r} = \mathbf{r}(\mathbf{x}, \mathbf{z}) = \begin{Bmatrix} \boldsymbol{\epsilon}^e - \boldsymbol{\epsilon}^{e \text{ tr}} + \Delta \lambda \partial_{\boldsymbol{\sigma}} Q \\ F(\bar{\boldsymbol{\sigma}}, \bar{p}_c) \end{Bmatrix}. \quad (38)$$

The solution of the problem is the vector \mathbf{x}^* such that $\mathbf{r}(\mathbf{x}^*) = \mathbf{0}$ for a given \mathbf{z} . However, the equation is nonlinear, so the solution \mathbf{x}^* must be determined iteratively. To this end, we use Newton’s method and evaluate the local Jacobian matrix

$$\mathbf{r}'(\mathbf{x})|_{\mathbf{z}} = \mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}. \quad (39)$$

Preserving the tensor notation for the matrices, we have

$$\mathbf{A}_{11} = \mathbf{I} + \Delta \lambda (\partial_{\boldsymbol{\sigma}\boldsymbol{\sigma}}^2 Q : \mathbf{c}^e + \partial_{\bar{p}_c}^2 Q \otimes \partial_{\boldsymbol{\epsilon}^e} \bar{p}_c) \quad (40)$$

$$\mathbf{A}_{12} = \Delta \lambda \partial_{\bar{p}_c}^2 Q \times \partial_{\Delta \lambda} \bar{p}_c + \partial_{\boldsymbol{\sigma}} Q \quad (41)$$

$$\mathbf{A}_{21} = \partial_{\boldsymbol{\sigma}} F : \mathbf{c}^e + \partial_{\bar{p}_c} F \times \partial_{\boldsymbol{\epsilon}^e} \bar{p}_c \quad (42)$$

$$\mathbf{A}_{22} = \partial_{\bar{p}_c} F \times \partial_{\Delta \lambda} \bar{p}_c, \quad (43)$$

where \mathbf{I} is the rank-four symmetric identity tensor with components $I_{ijkl} = (\delta_{ij}\delta_{kl} + \delta_{ij}\delta_{jk})/2$, $\mathbf{c}^e = \partial \bar{\boldsymbol{\sigma}} / \partial \boldsymbol{\epsilon}^e$ is the tangential elasticity tensor, and $(\partial_{\Delta \lambda}, \partial_{\boldsymbol{\epsilon}^e}) \bar{p}_c$ are the partial derivatives of \bar{p}_c obtained from the incremental hardening rule.

Since \mathbf{r} is zero at $\mathbf{x} = \mathbf{x}^*$, we can differentiate \mathbf{r} with respect to \mathbf{z} at $\mathbf{x} = \mathbf{x}^*$ to get

$$\frac{\partial \mathbf{r}}{\partial \mathbf{z}} = \frac{\partial \mathbf{r}}{\partial \mathbf{z}} \Big|_{\mathbf{x}} + \left(\frac{\partial \mathbf{r}}{\partial \mathbf{x}} \Big|_{\mathbf{z}} \right) \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \mathbf{0}, \quad (44)$$

which gives

$$\mathbf{A} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = - \frac{\partial \mathbf{r}}{\partial \mathbf{z}} \Big|_{\mathbf{x}} \Rightarrow \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = -\mathbf{B} \cdot \frac{\partial \mathbf{r}}{\partial \mathbf{z}} \Big|_{\mathbf{x}}, \quad (45)$$

where $\mathbf{B} = \mathbf{A}^{-1}$. The inverse exists provided that the local iteration has converged.

Expanding Eq. (45) gives

$$\begin{bmatrix} \partial_{\boldsymbol{\epsilon}^e} \boldsymbol{\epsilon}^e & \partial_s \boldsymbol{\epsilon}^e \\ \partial_{\Delta \lambda} \Delta \lambda & \partial_s \Delta \lambda \end{bmatrix} = - \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}, \quad (46)$$

where the the abbreviated expressions $\partial_{\boldsymbol{\epsilon}^e} \square = \partial \square / \partial \boldsymbol{\epsilon}^e \equiv \partial \square / \partial \boldsymbol{\epsilon}^{e \text{ tr}}$ and $\partial_s \square = \partial \square / \partial s \equiv \partial \square / \partial s^{\text{tr}}$ have been used. The submatrices are derived from the tensors

$$\mathbf{C}_{11} = \Delta \lambda \partial_{\bar{p}_c}^2 Q \otimes \partial_{\boldsymbol{\epsilon}^e} \bar{p}_c - \mathbf{I} \quad (47)$$

$$\mathbf{C}_{12} = \Delta \lambda \partial_{\bar{p}_c}^2 Q \times \partial_s \bar{p}_c \quad (48)$$

$$\mathbf{C}_{21} = \partial_{\bar{p}_c} F \times \partial_{\boldsymbol{\epsilon}^e} \bar{p}_c \quad (49)$$

$$\mathbf{C}_{22} = \partial_{\bar{p}_c} F \times \partial_s \bar{p}_c \quad (50)$$

We thus obtain

$$\boldsymbol{\alpha} = \frac{\partial \boldsymbol{\epsilon}^e}{\partial \boldsymbol{\epsilon}} = -\mathbf{B}_{11} : \mathbf{C}_{11} - \mathbf{B}_{12} \otimes \mathbf{C}_{21} \quad (51)$$

$$\boldsymbol{\beta} = \frac{\partial \boldsymbol{\epsilon}^e}{\partial s} = -\mathbf{B}_{11} : \mathbf{C}_{12} - \mathbf{B}_{12} \times \mathbf{C}_{22}. \quad (52)$$

Substituting the above expressions into (18) yields the tangent operators

$$\mathbf{c} = \frac{\partial \bar{\boldsymbol{\sigma}}}{\partial \boldsymbol{\epsilon}} = \mathbf{c}^e : \boldsymbol{\alpha}, \quad \mathbf{a} = -\frac{\partial \bar{\boldsymbol{\sigma}}}{\partial s} = -\mathbf{c}^e : \boldsymbol{\beta}. \quad (53)$$

The tensor \mathbf{a} above is the consistent derivative of the effective stress tensor with respect to capillary stress and accounts for full coupling of deformation and fluid flow. Note from the effective stress Eq. (4) that $\partial \bar{\boldsymbol{\sigma}}/\partial s \neq (\partial p^*/\partial s)\mathbf{1}$ even if $B = 1$, since $\partial \boldsymbol{\sigma}/\partial s \neq \mathbf{0}$ even if the applied external load is constant. The latter statement is analogous to the argument that the total stress tensor generally varies with time even if the applied external load remains fixed. This variation of the total stress with time is responsible for the Mandel-Cryer effect, which is a characteristic of the fully coupled solution, see Lambe and Whitman [50, p. 417]. To our knowledge, this is the first time that such consistent variation of the effective stress with respect to capillary pressure has been derived. Apart from its noteworthy physical significance, this derivative is also crucial for achieving the optimal convergence rate of Newton iteration.

4.2. Isotropy and spectral representation of tangent operators

By isotropy we mean that the constitutive model can be expressed in terms of the invariants of the stress tensor. It is important to note that strain localization is enhanced by the third stress invariant, so in this work we shall use all three invariants of the stress tensor in the constitutive formulation. For three-invariant models, spectral decomposition combined with return-mapping in principal elastic strain space provide an effective numerical scheme for stress-point integration [17,57,69].

To develop the relevant tangent operators we write the effective Cauchy stress tensor and the elastic strain tensor in spectral form

$$\bar{\boldsymbol{\sigma}} = \sum_{A=1}^3 \bar{\sigma}_A \mathbf{m}^{(A)}, \quad \boldsymbol{\epsilon}^e = \sum_{A=1}^3 \epsilon_A^e \mathbf{m}^{(A)}, \quad \boldsymbol{\epsilon}^{e \text{ tr}} = \sum_{A=1}^3 \epsilon_A^{e \text{ tr}} \mathbf{m}^{(A)}, \quad (54)$$

where $\mathbf{m}^{(A)} = \mathbf{n}^{(A)} \otimes \mathbf{n}^{(A)}$ is the spectral direction constructed from unit vector $\mathbf{n}^{(A)}$ in the direction of principal stress $\bar{\sigma}_A$. Note the coaxiality of the three tensors, particularly the coaxiality between $\boldsymbol{\epsilon}^e$ and $\boldsymbol{\epsilon}^{e \text{ tr}}$ that emanates from the fact that the plastic spin is zero in infinitesimal plasticity. This means that the spin of the principal axes is determined by the tensor $\boldsymbol{\epsilon}^{e \text{ tr}}$ alone.

We recall the following spectral form of \mathbf{c}^e :

$$\mathbf{c}^e = \sum_{A=1}^3 \sum_{B=1}^3 \mathcal{A}_{AB}^e \mathbf{m}^{(A)} \otimes \mathbf{m}^{(B)} + \frac{1}{2} \sum_{A=1}^3 \sum_{B \neq A}^3 \left(\frac{\bar{\sigma}_B - \bar{\sigma}_A}{\epsilon_B^e - \epsilon_A^e} \right) (\mathbf{m}^{(AB)} \otimes \mathbf{m}^{(AB)} + \mathbf{m}^{(AB)} \otimes \mathbf{m}^{(BA)}), \quad (55)$$

where $\mathbf{m}^{(AB)} = \mathbf{n}^{(A)} \otimes \mathbf{n}^{(B)}$ and \mathcal{A}_{AK}^e is the tangential elasticity matrix in principal axes. In a similar fashion, the spectral form of $\boldsymbol{\alpha} \equiv \partial \boldsymbol{\epsilon}^e / \partial \boldsymbol{\epsilon}$ is

$$\boldsymbol{\alpha} = \sum_{A=1}^3 \sum_{B=1}^3 \frac{\partial \epsilon_A^e}{\partial \epsilon_B} \mathbf{m}^{(A)} \otimes \mathbf{m}^{(B)} + \frac{1}{2} \sum_{A=1}^3 \sum_{B \neq A}^3 \left(\frac{\epsilon_B^e - \epsilon_A^e}{\epsilon_B^{e \text{ tr}} - \epsilon_A^{e \text{ tr}}} \right) (\mathbf{m}^{(AB)} \otimes \mathbf{m}^{(AB)} + \mathbf{m}^{(AB)} \otimes \mathbf{m}^{(BA)}), \quad (56)$$

Taking the inner product gives the spectral form of $\mathbf{c} = \mathbf{c}^e : \boldsymbol{\alpha}$ as

$$\mathbf{c} = \sum_{A=1}^3 \sum_{B=1}^3 \mathcal{A}_{AB} \mathbf{m}^{(A)} \otimes \mathbf{m}^{(B)} + \frac{1}{2} \sum_{A=1}^3 \sum_{B \neq A}^3 \left(\frac{\bar{\sigma}_B - \bar{\sigma}_A}{\epsilon_B^{e \text{ tr}} - \epsilon_A^{e \text{ tr}}} \right) (\mathbf{m}^{(AB)} \otimes \mathbf{m}^{(AB)} + \mathbf{m}^{(AB)} \otimes \mathbf{m}^{(BA)}), \quad (57)$$

where $\mathbf{m}^{(A)} = \mathbf{m}^{\text{tr}(A)}$ is the spectral direction of the elastic trial predictor strain, and

$$\mathcal{A}_{AB} = \frac{\partial \bar{\sigma}_A}{\partial \epsilon_B} = \sum_{K=1}^3 \frac{\partial \bar{\sigma}_A}{\partial \epsilon_K^e} \frac{\partial \epsilon_K^e}{\partial \epsilon_B} = \sum_{K=1}^3 \mathcal{A}_{AK}^e \frac{\partial \epsilon_K^e}{\partial \epsilon_B} \quad (58)$$

is the consistent tangential moduli matrix in principal axes.

We also recall the following spectral form for $\partial \bar{\boldsymbol{\sigma}}/\partial p$:

$$\frac{\partial \bar{\boldsymbol{\sigma}}}{\partial p} = \sum_{A=1}^3 \frac{\partial \bar{\sigma}_A}{\partial p} \mathbf{m}^{(A)} + \sum_{A=1}^3 \sum_{B \neq A}^3 \frac{\partial \Theta_{AB}}{\partial p} (\bar{\sigma}_B - \bar{\sigma}_A) \mathbf{m}^{(AB)}, \quad (59)$$

where Θ_{AB} is the rotation of principal axes. As noted before, this spin is determined by $\boldsymbol{\epsilon}^{e \text{ tr}}$ alone, and so $\partial \Theta_{AB}/\partial p \equiv 0$. Hence, we get the simplified relation

$$\frac{\partial \bar{\boldsymbol{\sigma}}}{\partial p} = \sum_{A=1}^3 \sum_{K=1}^3 \frac{\partial \bar{\sigma}_A}{\partial \epsilon_K^e} \frac{\partial \epsilon_K^e}{\partial p} \mathbf{m}^{(A)} = \sum_{A=1}^3 \sum_{K=1}^3 \mathcal{A}_{AK}^e \frac{\partial \epsilon_K^e}{\partial p} \mathbf{m}^{(A)}. \quad (60)$$

Note that the change occurs at fixed principal directions.

We now turn to obtaining the tangential matrix in principal axes. In this case, the local residual equation becomes

$$\mathbf{r}(\mathbf{x}, \mathbf{z}) = \begin{Bmatrix} \epsilon_1^e - \epsilon_1^{e \text{ tr}} + \Delta \lambda \partial_{\bar{\sigma}_1} Q \\ \epsilon_2^e - \epsilon_2^{e \text{ tr}} + \Delta \lambda \partial_{\bar{\sigma}_2} Q \\ \epsilon_3^e - \epsilon_3^{e \text{ tr}} + \Delta \lambda \partial_{\bar{\sigma}_3} Q \\ F(\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3, \bar{p}_c) \end{Bmatrix}, \quad \mathbf{x} = \begin{Bmatrix} \epsilon_1^e \\ \epsilon_2^e \\ \epsilon_3^e \\ \Delta \lambda \end{Bmatrix}, \quad \mathbf{z} = \begin{Bmatrix} \epsilon_1^{e \text{ tr}} \\ \epsilon_2^{e \text{ tr}} \\ \epsilon_3^{e \text{ tr}} \\ s^{\text{tr}} \end{Bmatrix}. \quad (61)$$

Iterating for the local solution \mathbf{x}^* corresponding to a given \mathbf{z} requires the evaluation of the 4×4 matrix,

$$\mathbf{r}'(\mathbf{x})|_{\mathbf{z}} = \mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{34} & A_{44} \end{bmatrix}, \quad (62)$$

where

$$A_{IJ} = \delta_{IJ} + \Delta \lambda \left(\sum_{K=1}^3 \frac{\partial^2 Q}{\partial \sigma_I \partial \sigma_K} \mathcal{A}_{KJ}^e + \frac{\partial^2 Q}{\partial \sigma_I \partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \epsilon_J^e} \right), \quad I, J = 1, 2, 3, \quad (63)$$

$$A_{I4} = \Delta \lambda \frac{\partial^2 Q}{\partial \sigma_I \partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \Delta \lambda} + \frac{\partial Q}{\partial \sigma_I}, \quad I = 1, 2, 3, \quad (64)$$

$$A_{4J} = \sum_{K=1}^3 \frac{\partial F}{\partial \sigma_K} \mathcal{A}_{KJ}^e + \frac{\partial F}{\partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \epsilon_J^e}, \quad J = 1, 2, 3 \quad (65)$$

and

$$A_{44} = \frac{\partial F}{\partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \Delta \lambda}. \quad (66)$$

As in the previous section we denote the inverse of \mathbf{A} by the 4×4 matrix \mathbf{B} with components $[B_{IJ}]$. Eq. (45)₂ then becomes

$$\frac{\partial \epsilon_I^e}{\partial \epsilon_J} = -\sum_{K=1}^4 B_{IK} C_{KJ}, \quad I, J = 1, 2, 3 \quad (67)$$

and

$$\frac{\partial \epsilon_I^e}{\partial s} = -\frac{\partial \epsilon_I^e}{\partial p} = -\sum_{K=1}^3 B_{IK} C_{K4}, \quad I = 1, 2, 3. \quad (68)$$

where

$$C_{IJ} = -\delta_{IJ} + \Delta \lambda \frac{\partial^2 Q}{\partial \bar{\sigma}_I \partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \epsilon_J^e}, \quad C_{I4} = \Delta \lambda \frac{\partial^2 Q}{\partial \bar{\sigma}_I \partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial s}, \quad (69)$$

$$C_{4J} = \frac{\partial F}{\partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial \epsilon_J^e}, \quad C_{44} = \frac{\partial F}{\partial \bar{p}_c} \frac{\partial \bar{p}_c}{\partial s}.$$

Note that the derivatives $\partial \Delta \lambda / \partial \epsilon_J$ and $\partial \Delta \lambda / \partial s$ are not used in the formulation.

It is important to recall that in evaluating $[A_{IJ}]$ the derivatives are evaluated with respect to \mathbf{x} with \mathbf{z} held fixed, whereas in eval-

uating $[C_{ij}]$ the derivatives are evaluated with respect to \mathbf{z} with \mathbf{x} held fixed.

4.3. Derivatives of \bar{p}_c

In critical state isotropic models, $-\bar{p}_c > 0$ is a measure of the size of the yield surface in the fully saturated state. For isotropic plasticity models, $-p_c$ is the distance from the origin of the stress space to the intersection of the compression cap with the hydrostatic axis. Experimental evidence [38] suggests that in the unsaturated regime the compression cap expands with increasing suction, and so $-\bar{p}_c$ must increase with increasing suction. An experimentally validated analytical form for \bar{p}_c reflecting this feature is given by the hardening law [14,38]

$$\bar{p}_c = -\exp[a(\xi)](-p_c)^{b(\xi)}, \tag{70}$$

where $p_c = p_c(\bar{\boldsymbol{\sigma}}, \Delta\lambda, \epsilon^{e\ tr})$ is the preconsolidation pressure in the fully saturated state, and $a(\xi)$ and $b(\xi)$ are functions of a so-called bonding variable ξ such that $a = 0, b = 1$, and $\bar{p}_c = p_c$ in the fully saturated state. The bonding variable ξ varies with the suction stress s through the equation

$$\xi = f(s)(1 - S_r), \quad f(s) = 1 + \frac{s/p_{atm}}{10.7 + 2.4(s/p_{atm})}. \tag{71}$$

The suction function $f(s)$ given above is a hyperbolic fit to Fisher’s [37] curve as suggested in [14]. For isothermal deformation the degree of saturation S_r may be expressed as a function of s alone, for example, through the van Genuchten relation [41]

$$S_r = S_1 + (S_2 - S_1) \left[1 + \left(\frac{s}{s_a} \right)^n \right]^{-m}, \tag{72}$$

where s_a is the air entry value and S_1, S_2, m , and n are fitting parameters. The relations presented above are highly nonlinear, but the derivatives can be obtained in a straightforward fashion.

Taking the derivative of \bar{p}_c with respect to a variable \square other than s gives

$$\frac{\partial \bar{p}_c}{\partial \square} = \exp(a)b(-p_c)^{b-1} \frac{\partial p_c}{\partial \square}, \tag{73}$$

where $\partial p_c / \partial \square$ can be evaluated from the specific form of the incremental hardening law. The derivative with respect to s itself is given by

$$\left. \frac{\partial \bar{p}_c}{\partial s} \right|_{\mathbf{x}} = -\exp(a)\xi'(s)(-p_c)^b [a'(\xi) + b'(\xi) \ln(-p_c)]. \tag{74}$$

We see that in both cases, s is separable from the other variables with respect to derivatives. This facilitates a straightforward implementation of any established critical state model within the framework of unsaturated poromechanics. We present two models below, one for clay and another for sand. In both models we assume an associated plastic flow in which the plastic potential function is the same as the yield function.

4.4. Three-invariant modified Cam-Clay

In the formulation of a three-invariant model, it is sometimes convenient to use the (\bar{p}, q, θ) representation analogous to the cylindrical Haigh–Westergaard coordinates [30], where the hydrostatic axis serves as the pole and any of the three positive (extensional) principal stress axes serves as the polar axis. We define

$$\bar{p} = \frac{1}{3} \text{tr}(\bar{\boldsymbol{\sigma}}), \quad q = \sqrt{\frac{3}{2}} \|\bar{\boldsymbol{\sigma}} - \bar{p}\mathbf{1}\|, \tag{75}$$

The polar radius extends to all polar directions $0 \leq \theta \leq 2\pi$, see Fig. 2. In addition, the ellipticity is defined as (see [77])

$$\rho = q_{\text{ext}}/q_{\text{com}}, \quad 1/2 \leq \rho \leq 1, \tag{76}$$

where $\sqrt{2/3}q_{\text{com}}$ (alternatively, $\sqrt{2/3}q_{\text{ext}}$) is the radius on the compressive (alternatively, extensional) principal stress axis. The ellipticity ρ describes the deviation from roundness of the cross section of the yield surface on the deviatoric plane. The upper bound $\rho = 1$ is for a circular cross section on the deviatoric plane, whereas the lower bound $\rho = 1/2$ is for a limiting triangular cross section.

A three-invariant modified Cam-Clay yield function is of the form

$$F = \zeta^2 \frac{q^2}{M^2} + \bar{p}(\bar{p} - \bar{p}_c) \leq 0, \tag{77}$$

where M is the slope of the critical state line (CSL) and $\zeta = \zeta(\rho, \theta)$ is a scaling function designed to reproduce the effect of ellipticity. The function ζ satisfies the following boundary conditions: (a) $\zeta = 1/\rho$ when $\theta = 0$; and (b) $\zeta = 1$ when $\theta = \pi/3$. Since θ does not depend on the first stress invariant, all cross sections of the yield surface are scaled similarly. Possible forms of ζ include those proposed by Willam and Warnke [77], Gudehus [45], and Argyris et al. [4] (the latter two have the same form). Fig. 3 shows the shape of the yield function on a compressional meridian plane (labeled MCC).

The hardening law at full saturation is given by the discrete evolution equation [14]

$$p_c = p_{c,n} \exp\left(\frac{\epsilon_v^e - \epsilon_v^{e\ tr}}{\lambda - \tilde{K}}\right), \tag{78}$$

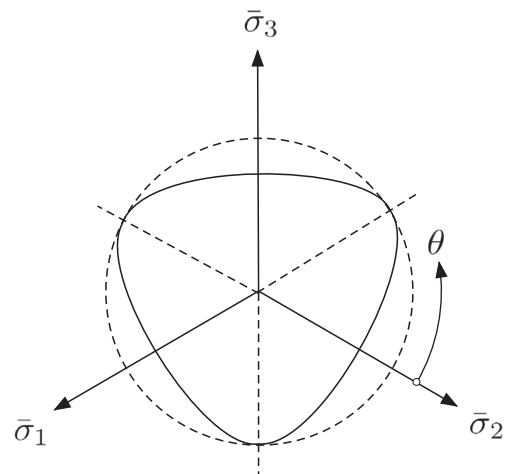


Fig. 2. Influence of the third stress invariant on the shape of the yield surface.

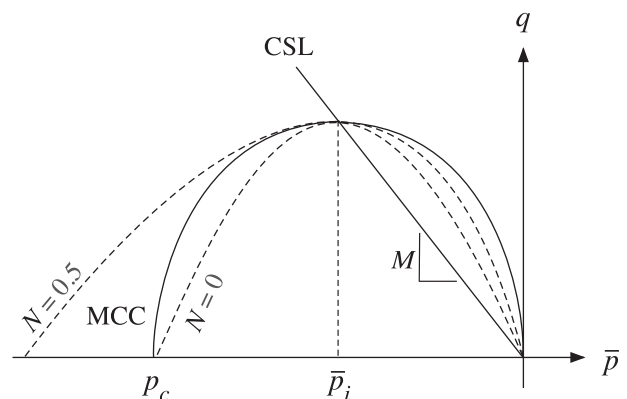


Fig. 3. Exponential ($N = 0, 0.5$) and elliptical (MCC) yield surfaces on compressional meridian plane. CSL = critical state line.

where $\epsilon_v^e = \text{tr}(\epsilon^e)$, $\epsilon_v^{e\text{tr}} = \text{tr}(\epsilon^{e\text{tr}})$, and $\tilde{\lambda}$ and $\tilde{\kappa}$ are compressibility parameters [14]. We recall that p_c is readily separable from suction with respect to differentiation, i.e., one can express $\bar{p}_c = \bar{p}_c(s, p_c)$. This specific evolution equation for p_c follows from $p_c = p_c(\bar{\sigma}(\epsilon^e), \epsilon^{e\text{tr}})$ and does not have $\Delta\lambda$ explicitly as one of its arguments.

4.5. Model for sand with state parameter

A three-invariant yield function for sand is given by [2,11,48]

$$F = \zeta q + \eta \bar{p} \leq 0, \quad (79)$$

where $\zeta = \zeta(\rho, \theta)$ is the same scaling function introduced in the previous section, and

$$\eta = \begin{cases} M[1 + \ln(\bar{p}_i/\bar{p})] & \text{if } N = 0; \\ (M/N)[1 - (1 - N)(\bar{p}/\bar{p}_i)^{N/(1-N)}] & \text{if } N > 0. \end{cases} \quad (80)$$

The parameter M has the same meaning as in the modified Cam-Clay model, and $\bar{p}_i < 0$ takes the role of the plastic internal variable, see Fig. 3. A closed-form expression for \bar{p}_i is obtained by setting $F = 0$,

$$\frac{\bar{p}_i}{\bar{p}} = \begin{cases} \exp(\eta/M - 1) & \text{if } N = 0; \\ [(1 - N)/(1 - \eta N/M)]^{(1-N)/N} & \text{if } N > 0. \end{cases} \quad (81)$$

The parameter N determines the curvature of the yield surface on the hydrostatic axis, and typically has a value less than 0.4 [48]. If $\zeta = 1$ the yield function reduces the original Cam-Clay yield function [67]. The form of the yield function readily provides expressions for p_c in terms of \bar{p}_i . Setting $\eta = 0$ in (82) and solving for \bar{p} gives the value of $p_c \equiv \bar{p}$

$$p_c = \bar{p}_i \times \begin{cases} e & \text{if } N = 0; \\ (1 - N)^{(N-1)/N} & \text{if } N > 0, \end{cases} \quad (82)$$

in which e is the natural logarithm constant.

4.6. Fluid flow derivatives

We consider a four-parameter soil–water characteristic curve derived from the van Genuchten [41] model of the form

$$S_r(p) = S_1 + (S_2 - S_1) \left[1 + \left(\frac{s}{s_a} \right)^{n-m} \right]^{-m}, \quad (83)$$

where $s = -p$, S_1 is the residual water saturation, S_2 is the maximum water saturation, s_a is a scaling pressure, and n and m are empirical constants defining the shape of the saturation curve. The constants n and m are not independent, but are related as $m = (n - 1)/n$. Thus, the model has a total of four independent parameters. The relative permeability of the water phase is similarly defined as

$$k_{rw}(\theta) = \theta^{1/2} \left[1 - \left(1 - \theta^{1/m} \right)^m \right]^2, \quad \theta = \frac{S_r(p) - S_1}{S_2 - S_1}. \quad (84)$$

The pressure derivative of saturation can be readily evaluated as

$$S'_r(p) = (S_2 - S_1) \frac{n-1}{s_a} \left(\frac{s}{s_a} \right)^{n-1} \left[1 + \left(\frac{s}{s_a} \right)^n \right]^{-(1+m)}. \quad (85)$$

The derivative of the relative permeability is then given by

$$k'_{rw}(p) = k'_{rw}(\theta) \frac{S'_r(p)}{S_2 - S_1}, \quad (86)$$

where

$$k'_{rw}(\theta) = 2\theta^{(1/m-1/2)}(1 - \theta^{1/m})^{m-1} \left[1 - \left(1 - \theta^{1/m} \right)^m \right] + \frac{1}{2}\theta^{-1/2} \left[1 - \left(1 - \theta^{1/m} \right)^m \right]^2. \quad (87)$$

The above constitutive relations do not explicitly account for the effect of tortuosity and pore shape on the relative hydraulic conductivity (see [42,80] for more in-depth discussions of these subjects).

5. Numerical examples

In the three examples below, we simulate plane strain compression on rectangular specimens under globally undrained but locally drained conditions. This means that fluid can migrate inside the sample but is not free to enter or leave through the exterior boundaries of the sample. In the first example, a soil with a uniform density but with a spatially varying degree of saturation is modeled with the three-invariant modified Cam-Clay theory. The second example presents a similar simulation but uses the constitutive model for sand. The third example simulates spatially varying density and degree of saturation with the three-invariant plasticity theory for sand. Note that the three-invariant modified Cam-Clay theory has no state parameter and cannot account for the effect of initial density variation, so it cannot accommodate the conditions of the third example. Throughout the simulations, we use stabilized low-order (bilinear) quadrilateral elements with polynomial pressure projection stabilization [75,76] to suppress spurious pore pressure oscillations in the incompressible and nearly incompressible regimes.

5.1. Triggering a shear band in clay

The material parameters for the clay model are summarized in Tables 1 and 2. The values of the parameters are similar to those considered in [14]. The problem domain is a rectangular specimen 5 cm wide and 10 cm tall modeled with 200 stabilized four-node quadrilateral mixed elements with displacement and pressure degrees of freedom at each node. The block is supported on vertical rollers at the top and bottom boundaries (except at a bottom corner node that is pinned for stability), and compressed vertically at a rate of 0.001 cm/s so that in the absence of any non-uniform field the deformation would be homogeneous. Therefore, any calculated inhomogeneous deformation can be attributed directly to the initial field condition, which, in this case, is the spatial variation of the degree of saturation. The total simulation time is 350 s over 117 load steps.

The initial degree of saturation is shown in Fig. 4a and has been randomly generated to represent the effect of sample preparation in which soils are deposited horizontally in thin layers and sprayed unevenly with water before depositing the next horizontal layers. The result is a saturation distribution characterized by intermittent patches of wet and dry layers. The two vertical boundaries of the block are subjected to a pressure of 100 kPa that is held constant throughout the simulation. Since the degree of saturation varies throughout the block, the capillary pressure and the effective stresses also vary, creating an initially inhomogeneous stress field. The variable saturation also creates a pressure gradient field that induces fluid migration inside the block. The fluid flow is initially er-

Table 1
Solid deformation parameters for unsaturated clay.

Symbol	Value	Parameter
$\tilde{\kappa}$	0.03	Elastic compressibility
p_0	-0.1 MPa	Reference pressure
ϵ_{r0}^e	0.0	Reference strain
μ_0	10 MPa	Shear modulus
M	1.2	Critical state parameter
$\tilde{\lambda}$	0.09	Plastic compressibility
v_{c0}	1.95	Reference specific volume
ρ	7/9	Ellipticity

Table 2
Fluid conduction parameters for unsaturated clay.

Symbol	Value	Parameter
k	1.0×10^{-5} cm/s	Saturated hydraulic conductivity
ψ_1	0.0	Water retention parameter
ψ_2	1.0	Water retention parameter
n	2.0	Water retention parameter
s_a	0.01 MPa	Air entry pressure
c_1	0.185	Parameter of Ref. [38]
c_2	1.49	Parameter of Ref. [38]
p_{atm}	101.3 kPa	Atmospheric pressure

ratio owing to the initially erratic variation of degree of saturation, but eventually the gradient field smoothens out as shown in Figs. 4d.

Establishing a statically admissible initial condition is important for interpretation of the results. Prior to the first load step, initial effective stresses at the Gauss points are tentatively prescribed to balance the externally applied pressure of 100 kPa. However, randomly generated capillary pressures are also prescribed at the nodes. In general, the effective stresses and pore pressures will not balance the applied external pressure, and there will be some residual nodal forces. The first load step is then used to iteratively balance these forces. In the simulations presented below, we take the ‘initial condition’ as the converged solution after the first load step. The contour of saturation shown in Fig. 4a is the statically admissible saturation configuration after the first load step.

Since the degree of saturation varies within the problem domain, both the capillary pressure and the effective stresses also vary within the domain. The constitutive model for clay assumes the same critical state line for all the elements, so for the same reference specific volume the ‘initial condition’ is characterized by a nonuniform density field. However, this density field is not specified, but instead is calculated by the constitutive model from the prescribed initial degree of saturation and applied forces. In contrast, the density field and degree of saturation are two separately specified state variables for the constitutive model for sand.

Fig. 5 indicates a complex evolution of the stress ratio $-q/\bar{p}$ occurring within the problem domain as the specimen is compressed vertically and the fluid migrates locally within the specimen. Starting from the isotropic stress state in Fig. 5a, the effect of variable saturation is ‘felt’ almost instantaneously by the solid matrix as soon as the compression commences, as shown in Fig. 5b. Further vertical compression leads to more complicated stress pattern portrayed in Fig. 5c, where some points even reached the critical state. The final stress state, shown in Fig. 5d, suggests a stress ratio value of around 0.6 prevailing within and around the deformation band, which is well below the critical value of $M = 1.2$. At no time in the solution did the stress ratio exceed the critical value M , indicating that plasticity is restricted to the compression side of the yield surface.

The resulting deformation field for the clay simulation is shown in Fig. 6 and suggests the formation of a compactive shear band, a type of deformation band where shearing is accompanied by a reduction in volume over a narrow zone [15,16]. The volume reduction is due to the compaction of the air voids within the band that increases the degree of saturation as shown in Fig. 4. As the degree of saturation increases the pore water pressure increases, producing a pressure gradient field that enhances fluid migration away from the band. These complex multiphysical processes are illustrated further by the fluid flow vectors shown in Fig. 7a, suggesting that fluid is continually squeezed out of the band as the domain is compressed.

Fig. 7b shows the contour of the determinant of the drained acoustic tensor [14,64] at the end of the simulation. This determinant is an indicator of the propensity of a material to form a comp-

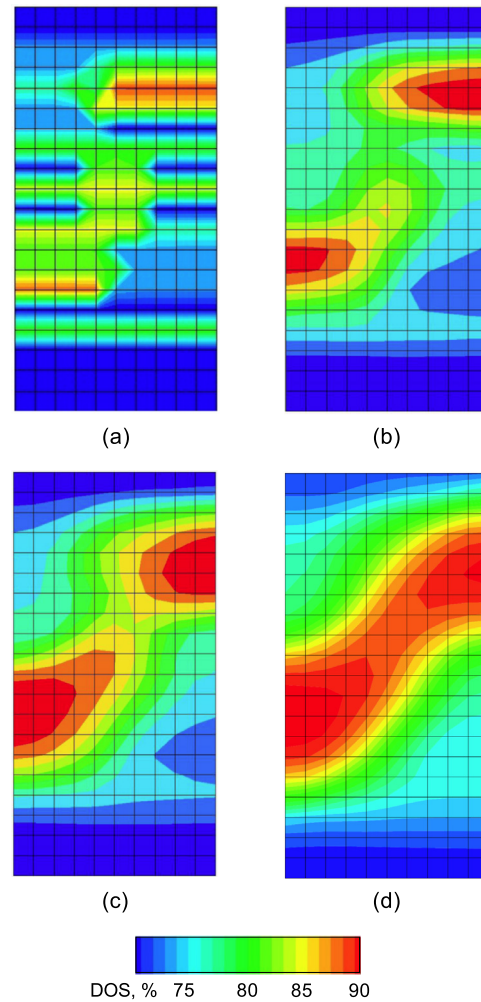


Fig. 4. Evolution of degree of saturation (DOS) in a partially saturated clay during vertical compression: (a) initial condition; and at axial strains of (b) 1.0%, (c) 2.0%, and (d) 3.5%.

active shear band under a locally drained condition [13,24]. The figure clearly shows a tendency to form a deformation band in the region where the determinant of the acoustic tensor vanishes. However, this band is fairly diffuse and the determinant function never reaches the value zero. This is because the stress point remains on the compression cap of the yield surface throughout the simulation. The compression cap is considered to be a stable region where the plastic modulus is always positive; hence, localized bifurcation in which the determinant becomes zero is not reached.

Fig. 8 shows the global convergence of Newton iteration for the unsaturated clay simulation. The residual vector is a composite vector consisting of out-of-balance nodal forces and fluid fluxes, and has been normalized with respect to its initial L_2 -norm. The figure shows that convergence of the iterations is asymptotically quadratic throughout the simulations. This implies that both the solid deformation and fluid flow equations have been consistently linearized. For the record, Step #117 is the last load increment in the simulation.

5.2. Triggering a shear band in sand

Tables 3 and 4 summarize the relevant material parameters for the sand model. The values of the parameters are similar to those

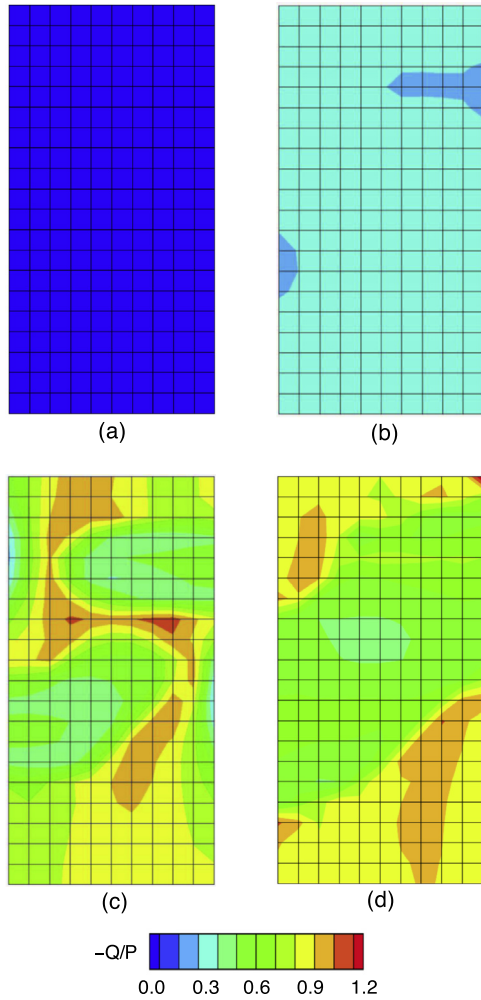


Fig. 5. Evolution of stress ratio $-q/\bar{p}$: (a) initial isotropic condition; and at axial strains of (b) 0.3%, (c) 1.1%, and (d) 3.5%. Note: red in the color bar is critical state, $-q/\bar{p} = M$.

considered in [11]. The finite element mesh and boundary conditions are the same as those used in the previous example. Note that the saturated hydraulic conductivity is two orders of magnitude higher than in the clay example. To capture a comparable range of deformation, the sample is deformed at a higher speed of 0.1 cm/s, which is equivalent to applying a total vertical compression of 0.45 cm over a period of 4.5 s. The total vertical compression is applied in 225 increments.

Fig. 9 shows the degree of saturation at two different stages of vertical compression. Once again, the zone of higher saturation concentrates in the neighborhood of the band where the soil undergoes greater compaction and shearing, see Fig. 10. Compaction of the air void is responsible for the increase in saturation, as can be seen by comparing Figs. 9 and 10. Note a strong correlation between the zone of greatest compaction with the zone of highest degree of saturation. Fig. 11a shows that fluid is expelled from the band as the compaction of this zone takes place. The determinant of the acoustic tensor in Fig. 11b shows that the constitutive model for sand gives rise to more pronounced localized deformation as the determinant function switches in sign within a narrow band. This implies an impending shear band forming, in contrast to the more diffuse deformation pattern predicted by the clay model that does not produce a reversal in the sign of the determinant function. Once again, Fig. 12 shows that convergence of the global Newton iteration is rapid.

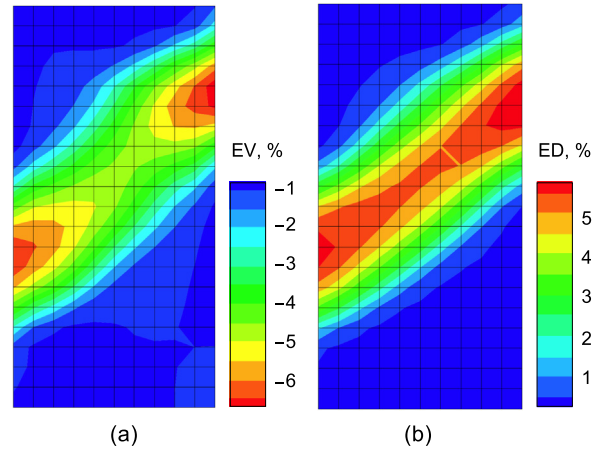


Fig. 6. Localized deformation in clay after applying a nominal axial compression of 3.5%: (a) volumetric strain, (b) deviatoric strain.

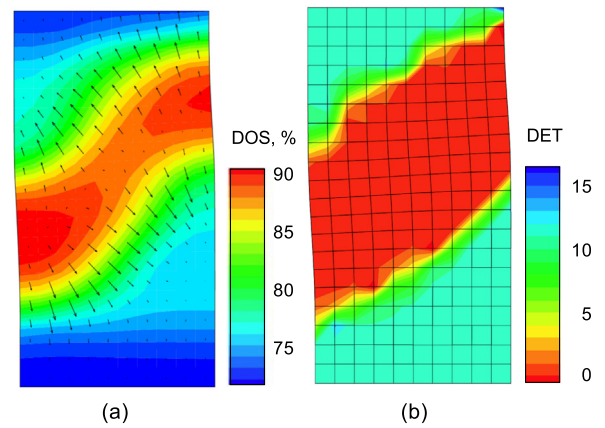


Fig. 7. Deformed mesh in clay at axial strain of 3.5%: (a) degree of saturation with fluid flow vectors, (b) normalized determinant of drained acoustic tensor.

5.3. Shear band in sand with variable density and saturation

The third example deals with spatially varying density and degree of saturation in sand. The density variation is derived from a digitally processed CT image of sand with specific volume varying from 1.3 to 2.0 (see Ref. [7] for details of the laboratory test). This

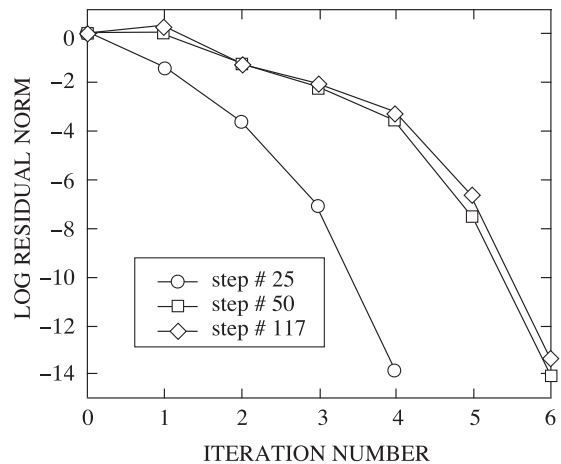


Fig. 8. Global convergence of unsaturated three-invariant clay model.

Table 3
Solid deformation material parameters for unsaturated sand.

Symbol	Value	Parameter
$\tilde{\kappa}$	0.03	Compressibility
p_0	-0.1 MPa	Reference pressure
e_{v0}^e	0.0	Reference strain
μ_0	20 MPa	Shear modulus
M	1.2	Critical state parameter
$\tilde{\lambda}$	0.11	Compressibility parameter
N	0.4	Yield surface parameter
h	280	Hardening modulus
v_{c0}	1.95	Reference specific volume
ρ	7/9	Ellipticity
α	-3.5	Limit dilatancy parameter
ρ_s	2.0 Mg/m ³	Solid density
ρ_w	1.0 Mg/m ³	Fluid density

Table 4
Fluid flow material parameters for unsaturated sand.

Symbol	Value	Parameter
k	1.5×10^{-3} cm/s	Saturated hydraulic conductivity
ψ_1	0.0	Water retention parameter
ψ_2	1.0	Water retention parameter
s_a	0.01 MPa	Air entry value of bubbling pressure
n	2.0	Constant in Von Genuchten equation
c_1	0.185	Parameter of Ref. [38]
c_2	1.49	Parameter of Ref. [38]
p_{atm}	101.3 kPa	Atmospheric pressure

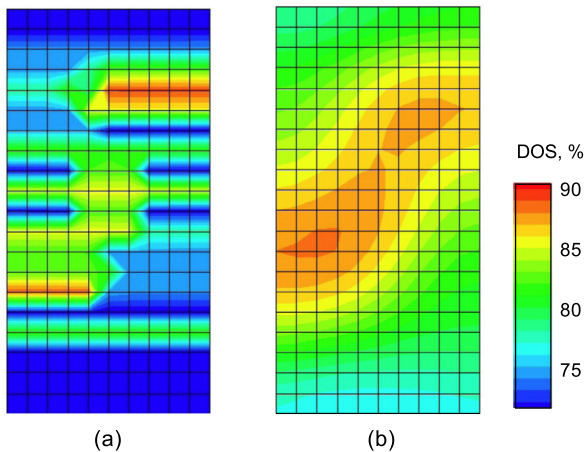


Fig. 9. Degree of saturation on unsaturated fine sand before and after vertical compression: (a) initial condition, (b) condition at nominal axial compression of 4.5%.

density contrast in the physical specimen is very strong and could very well dominate the formation of persistent shear band, so for the simulations we consider a similar density variation but with a smaller density contrast. The specimen is 137 mm tall, 39.5 mm wide, and 79.7 mm deep (out-of-plane), with specific volume varying from 1.4 to 1.8, see Fig. 13. The top and bottom boundaries are supported on vertical rollers that permit unconstrained shear band propagation, mimicking the conditions for the specimen described in Ref. [7]. On the lower half of the specimen is a locally loose layer. Preliminary simulations on this specimen assuming dry condition and using the same material parameters as in Example #2 indicate that when the specimen is compressed vertically, the loose layer simply compacts with no shear band forming in the specimen (Figs. 14). In the following simulations we show that the presence of moisture can trigger a shear band.

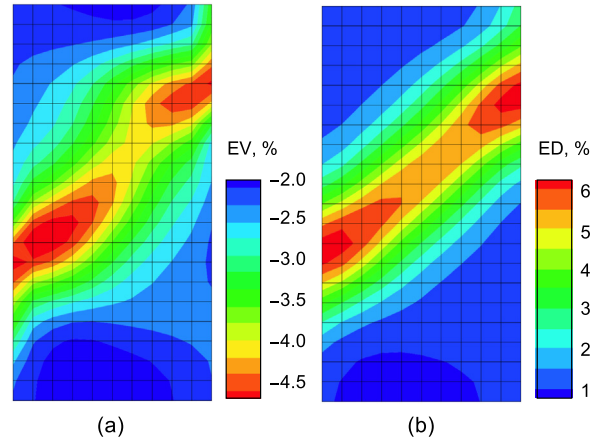


Fig. 10. Localized deformation in sand after applying a nominal axial compression of 4.5%: (a) volumetric strain, (b) deviatoric strain.

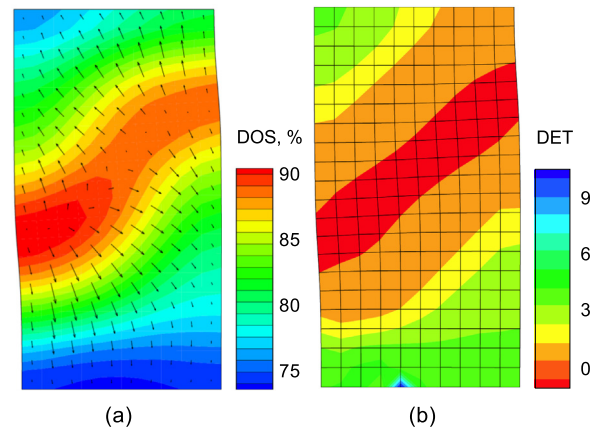


Fig. 11. Deformed mesh in sand after applying a nominal axial compression of 4.5%. Contours represent: (a) degree of saturation with fluid flow vectors, (b) normalized determinant of drained acoustic tensor.

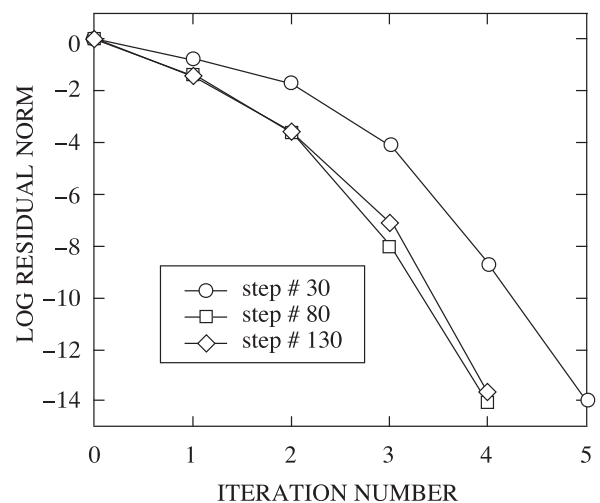


Fig. 12. Global convergence of unsaturated three-invariant sand formulation.

Furthermore, we show that the position and orientation of this band depend on the spatial distribution of the initial degree of saturation.

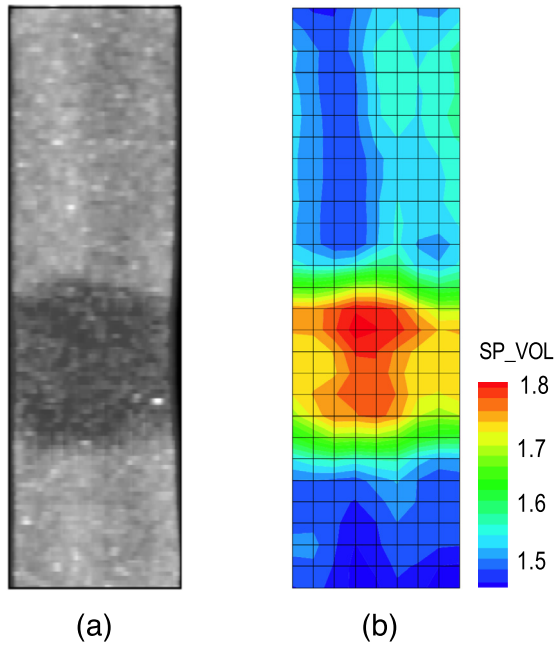


Fig. 13. Dry silica-concrete sand subjected to plane strain compression: (a) CT image with specific volume varying from 1.3 to 2.0 [7]; and (b) similar density variation but with specific volume adjusted to vary from 1.4 to 1.8.

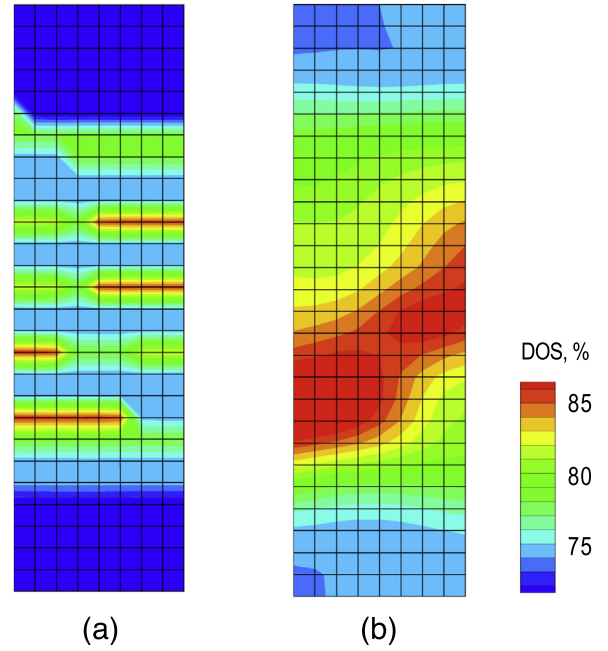


Fig. 15. Case #1: Degree of saturation (DOS) for partially saturated silica-concrete sand specimen subjected to vertical compression in plane strain: (a) initial condition; (b) condition at nominal vertical strain of 2.8%.

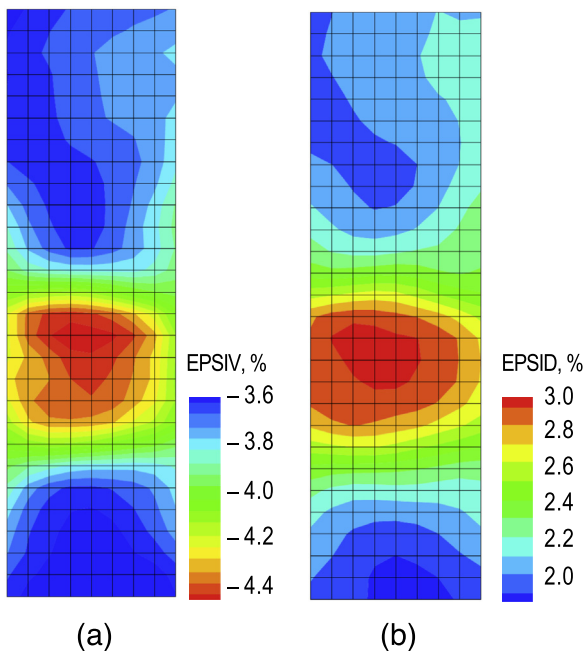


Fig. 14. Dry silica-concrete sand subjected to 4.5% vertical compression in plane strain: (a) volumetric strain; and (b) second invariant of deviatoric strain.

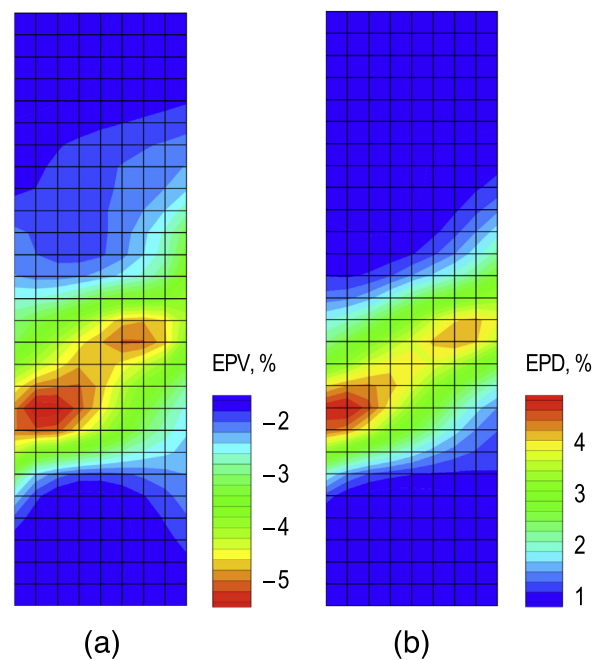


Fig. 16. Case #1: Volumetric (EPV) and deviatoric (EPD) strains in the specimen after a nominal vertical compression of 2.8%.

We consider two randomly generated saturation profiles superimposed on the sample with imposed initial heterogeneity in density. The first is shown in Fig. 15a and resembles the saturation profile considered in Example #2. The soil is assumed to have the same hydrological parameters as in Example #2. The specimen is compressed vertically at the rate of 0.002 cm/s until a persistent shear band can be observed. Fig. 15b shows the degree of saturation profile at 2.8% nominal vertical strain suggesting a trend toward full saturation within a narrow inclined region that ascends

in the rightward direction. As in the previous examples, the loose layer compacts as the shear band forms, causing the air voids to decrease and the degree of saturation to increase. Fig. 16 affirms the compaction-shearing deformation pattern occurring within the band. The flow gradient induces fluid migration away from the band, as indicated by the flow vectors shown in Fig. 17a. The determinant function shown in Fig. 17b indicates that the localization function changes sign, suggesting the formation of a persistent shear band. Fig. 18 demonstrates that the global convergence of Newton iteration for this particular simulation remains strong.

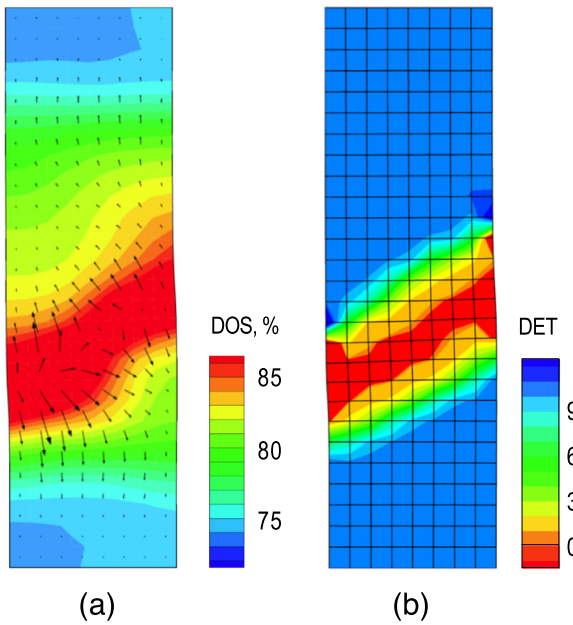


Fig. 17. Case #1: (a) Flow vectors superimposed with degree of saturation; and (b) normalized determinant function superimposed on deformed meshes. Snapshots taken after a nominal vertical compression of 2.8%.

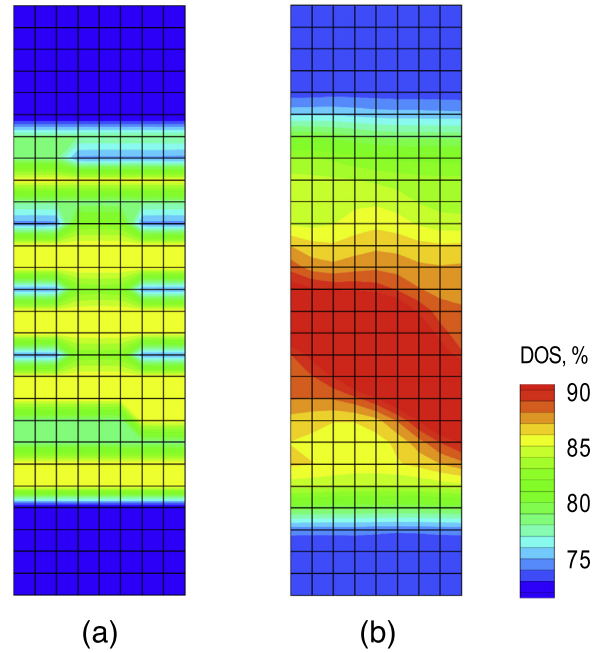


Fig. 19. Case #2: Degree of saturation (DOS) for partially saturated silica-concrete sand specimen subjected to vertical compression in plane strain: (a) initial condition; (b) condition at nominal vertical strain of 2.4%.

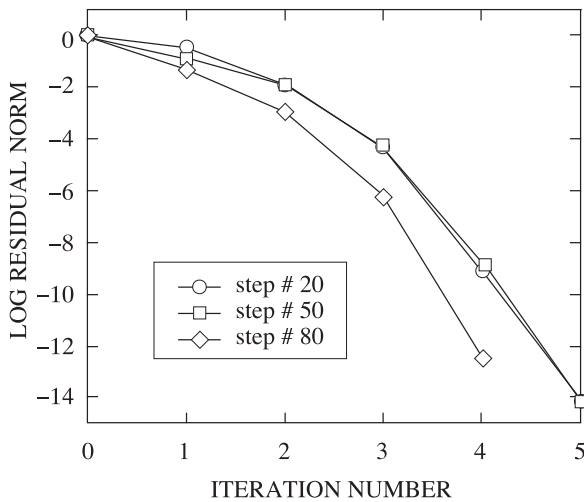


Fig. 18. Case #1: Global convergence of unsaturated three-invariant sand formulation with spatially varying density and degree of saturation.

Next, we consider a second sample with a randomly generated degree of saturation profile shown in Fig. 19a and with the same heterogeneous density distribution as in the first sample. After compressing the sample vertically to a nominal vertical strain of 2.4%, a nearly saturated band forms, but this time the band descends in the rightward direction as shown in Fig. 19b. This orientation is conjugate to the shear band in the previous example, demonstrating that the local saturation does have impact on the orientation of the shear band. The strain contours of Fig. 20 suggest that the pattern of localized deformation is dominated by combined compaction and shearing. As the deformation band compacts, fluid migrates into the surrounding zone (Fig. 21a). This is accompanied by the sign of the localization function reversing within the band, indicating localized bifurcation (Fig. 21b). Once again, the global convergence of Newton iterations remains strong (Fig. 22).

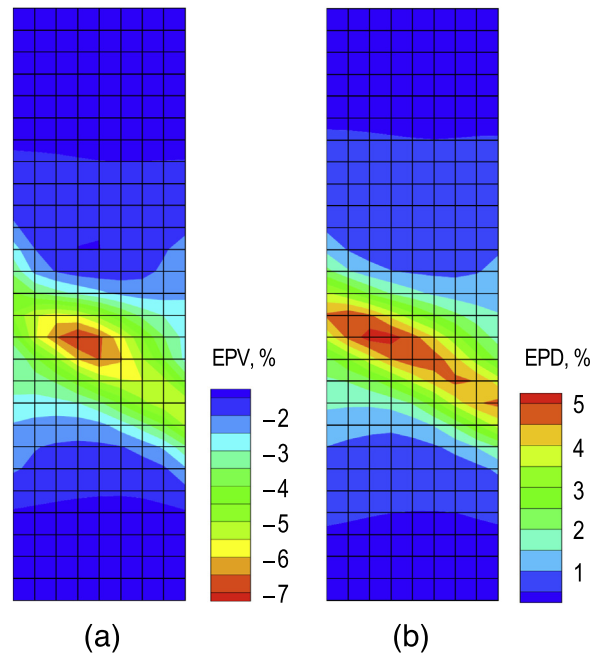


Fig. 20. Case #2: (a) Volumetric strain (EPV); and (b) deviatoric (EPD) strain. Snapshots taken after a nominal vertical compression of 2.4%.

Conducting a mesh sensitivity study is not straightforward for boundary-value problems with imposed material heterogeneities, because as the mesh is refined the description of material heterogeneity must also be refined. Obviously, mesh refinement is limited by the particulate nature of granular materials and should not go beyond the representative elementary volume. As noted in the Introduction, the heterogeneity is typically quantified from digital processing of a CT image, so the mesh refinement is limited

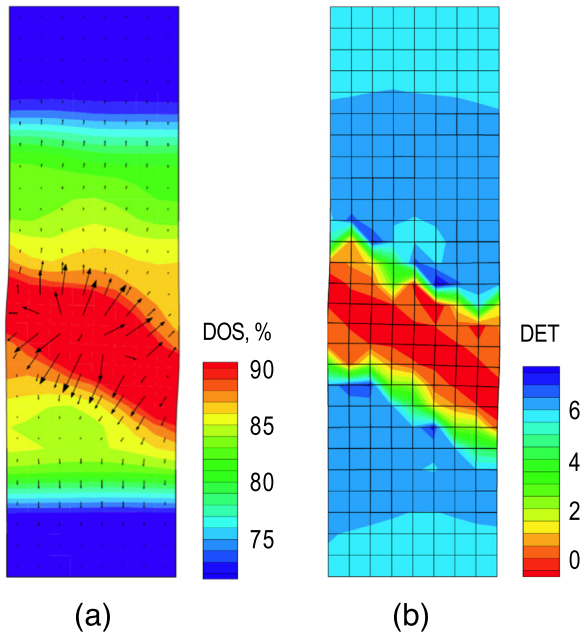


Fig. 21. Case #2: (a) Flow vectors superimposed with degree of saturation; and (b) normalized determinant function. Snapshots on deformed meshes taken after a nominal vertical compression of 2.4%.

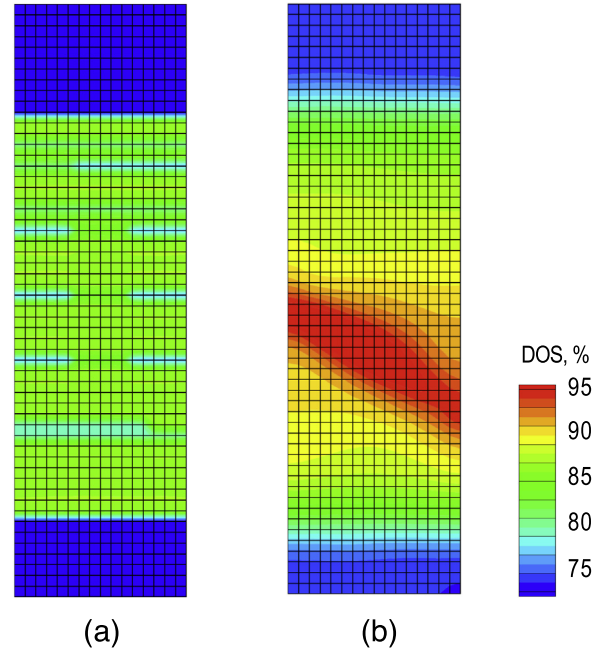


Fig. 23. Case #2-refined mesh: Degree of saturation (DOS) for partially saturated silica-concrete sand specimen subjected to vertical compression in plane strain: (a) initial condition; (b) condition at nominal vertical strain of 2.4%.

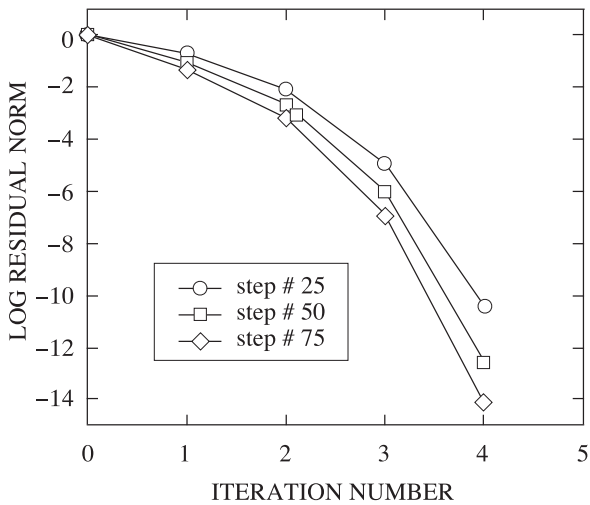


Fig. 22. Case #2: Global convergence of unsaturated three-invariant sand formulation with spatially varying density and degree of saturation.

by the resolution of the CT image. In the following, we consider a soil sample with similar density and saturation variations as in the previous example (Case #2).

The mesh consists of 16×55 stabilized mixed elements with an initial saturation variation similar to Case #2, as shown in Fig. 23a. After applying a nominal vertical strain of 2.4%, a similar compactive shear band emerges, i.e., descending to the right. The reduction of air voids within the band results in increased saturation (Fig. 23b). The pattern of persistent shear band characterized by significant volumetric and deviatoric strains is depicted in Fig. 24. In general, this pattern is similar to Case #2 but with a much better resolution. The pattern of fluid flow is also similar, as is the reversal in sign of the localization function (Fig. 25). The positions and inclinations of the shear bands predicted by the coarser and finer mesh simulations are essentially the same.

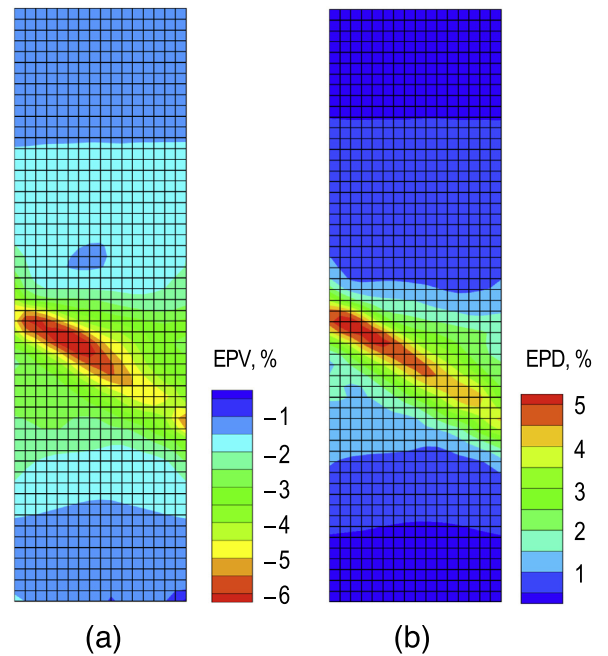


Fig. 24. Case #2-refined mesh: (a) Volumetric (EPV) strain; and (b) deviatoric (EPD) strain. Snapshots taken after a nominal vertical compression of 2.4%.

Fig. 26 compares the load versus displacement responses generated by the coarser and finer meshes for Case #2. The vertical load represents the resultant force at the prescribed vertical compression of the sample, and is calculated from the nodal pore pressures and effective stresses at the Gauss points projected to the upper boundary nodes. Prior to the peak load the two curves compare well, with the finer mesh exhibiting the expected slightly softer response compared to the coarser mesh. However, the two curves diverge beyond the peak loads, suggesting mesh sensitivity afflicting the two solutions. Bifurcation has been detected prior to the peak loads, with

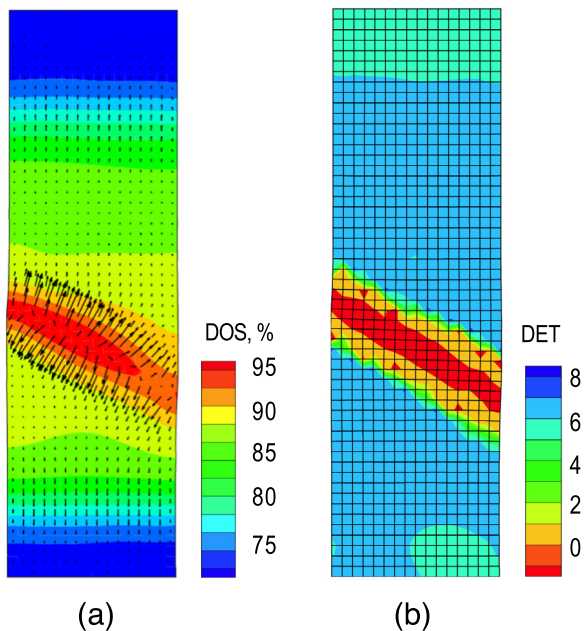


Fig. 25. Case #2-refined mesh: (a) Flow vectors superimposed with degree of saturation; and (b) normalized determinant function. Snapshots on deformed meshes taken after a nominal vertical compression of 2.4%.

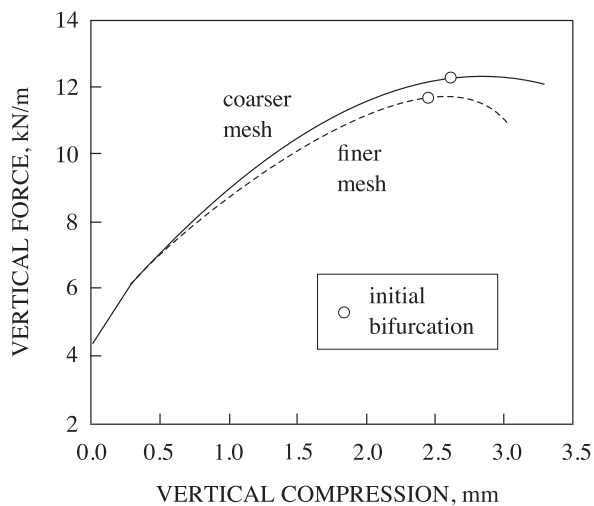


Fig. 26. Mesh sensitivity for Case #2: Prior to the peak load, the calculated responses compare well, with the finer mesh exhibiting a slightly softer response, a typical result. Beyond the peak points, the two solutions exhibit mesh sensitivity.

the finer mesh showing the expected propensity for an earlier bifurcation. It has been argued in [7] that for simulations involving strongly heterogeneous fields (e.g., strong density contrast), post-localization finite element enhancements should be introduced only after the persistent shear band has been fully identified, and not at the first onset of bifurcation. In the present case, the first onset of bifurcation as well as the peak load occurred almost at the same time as the formation of a persistent shear band, so the decision on when to introduce the post-localization enhancements is quite straightforward. Fig. 27 shows that the rate of convergence of Newton iteration remains rapid with the more refined mesh.

5.4. General remarks on the pattern of persistent shear band

The orientation of the persistent shear band described in this paper must not be confused with the analytical solution for a

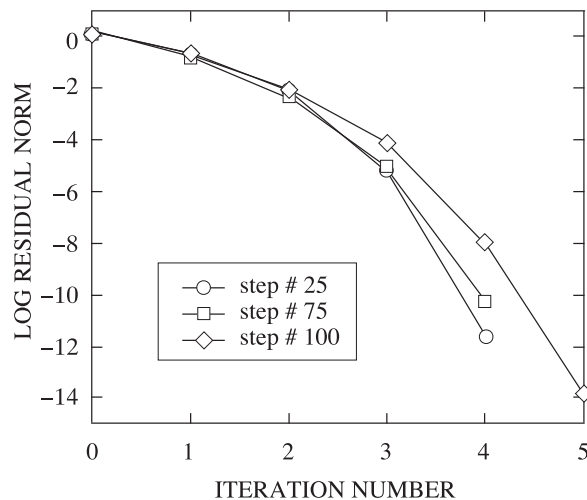


Fig. 27. Case #2: Global convergence of unsaturated three-invariant sand formulation with spatially varying density and degree of saturation: refined mesh.

homogeneous stress state determined by classic bifurcation analysis [64]. The persistent shear band reflects a general trend for a heterogeneous structure and not for a homogeneously stressed point. For the numerical examples discussed in this paper, the persistent shear band reflects the collective effect of compaction-induced deformation taking place in some regions, and dilation-induced shearing occurring in other regions. In the presence of fluid flow, the orientation of the persistent shear band also depends on the deformation rate due to the volume constraint it imposes on the overall deformation. Moreover, the persistent shear band also depends on some other random fields that cannot be fully quantified deterministically [2,29,46,59,68,70].

When the bifurcation condition is met at a particular Gauss point, a common procedure is to enhance the first bifurcating element to accommodate a post-localization mode, and then trace the evolution of the band [19,18,61]. However, as discussed in [7], this procedure will create unwanted bias in the presence of spatially distributed heterogeneous fields (such as density and/or degree of saturation). Some Gauss points could undergo bifurcation in different places, only to regain stability as the solution resolves the persistent shear band. In the presence of spatially distributed heterogeneous fields, it seems expedient to introduce the post-bifurcation enhancements after the persistent shear band has been fully developed. Once the persistent shear band has been identified, the post-localization enhancements are fairly straightforward.

6. Summary and conclusions

We have presented a mathematical framework for triggering a shear band in unsaturated granular materials with spatially varying degree of saturation and density. Both density and saturation have first-order effects on the persistent shear band. The volume constraint imposed by the presence of moisture, even in the unsaturated state, enhances the development of shear band. We have also presented a closed-form expression for the variation of the effective stress tensor with capillary pressure accounting for full coupling of the solid deformation and fluid flow. The performance of Newton iteration has been optimal in all the examples presented, demonstrating that this robust iterative technique can be applied successfully to some of the most challenging problems in geomechanics.

Acknowledgments

The authors are grateful to the three anonymous reviewers for their constructive reviews. Their comments were truly helpful in improving the paper from its first version. The authors are also grateful to Professor Rolando P. Orense for providing the source file for Fig. 1. Support for this work was provided by the US National Science Foundation (NSF) under Contract Numbers CMS-0324674 and CMMI-0936421 to Stanford University, and by Fonds zur Förderung der wissenschaftlichen Forschung (FWF) of Austria under Project Number L656-N22 to Universität für Bodenkultur.

References

- [1] P. Aeby, J. Forrer, C. Steinmeier, H. Fluhler, Image analysis for determination of dye tracer concentrations in sand columns, *Soil Sci. Am. J.* 61 (1997) 33–35.
- [2] J.E. Andrade, R.I. Borja, Capturing strain localization in dense sands with random density, *Int. J. Numer. Meth. Engrg.* 67 (2006) 1531–1564.
- [3] J.E. Andrade, R.I. Borja, Quantifying sensitivity of local site response models to statistical variations in soil properties, *Acta Geotech.* 1 (2006) 3–14.
- [4] J.H. Argyris, G. Faust, J. Szimmat, E.P. Warnke, K.J. Willam, Recent developments in the finite element analysis of prestressed concrete reactor vessels, *Nucl. Engrg. Des.* 28 (1974) 42–75.
- [5] P.B. Bochev, C.R. Dohrmann, M.D. Gunzburger, Stabilization of low-order mixed finite elements for the Stokes equations, *SIAM J. Numer. Anal.* 44 (2006) 82–101.
- [6] P.B. Bochev, C.R. Dohrmann, A computational study of stabilized low-order C^0 finite element approximations of Darcy equations, *Comput. Mech.* 38 (2006) 323–333.
- [7] R.I. Borja, X. Song, A. Rechenmacher, S. Abedi, W. Wu, Shear band in sand with spatially varying density, *J. Mech. Phys. Solids* 61 (2013) 219–234.
- [8] R.I. Borja, X. Liu, J.A. White, Multiphysics hillslope processes triggering landslides, *Acta Geotech.* 7 (2012) 261–269.
- [9] R.I. Borja, J.A. White, X. Liu, W. Wu, Factor of safety in a partially saturated slope inferred from hydro-mechanical continuum modeling, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 236–248.
- [10] R.I. Borja, A. Koliji, On the effective stress in unsaturated porous continua with double porosity, *J. Mech. Phys. Solids* 57 (2009) 1182–1193.
- [11] R.I. Borja, J.E. Andrade, Critical state plasticity. Part VI: Meso-scale finite element simulation of strain localization in discrete granular materials, *Comput. Meth. Appl. Mech. Engrg.* 195 (2006) 5115–5140.
- [12] R.I. Borja, On the mechanical energy and effective stress in saturated and unsaturated porous continua, *Int. J. Solids Struct.* 43 (2006) 1764–1786.
- [13] R.I. Borja, Condition for liquefaction instability in fluid-saturated granular soils, *Acta Geotech.* 1 (2006) 211–224.
- [14] R.I. Borja, Cam-Clay plasticity. Part V: a mathematical framework for three-phase deformation and strain localization analyzes of partially saturated porous media, *Comput. Meth. Appl. Mech. Engrg.* 193 (2004) 5301–5338.
- [15] R.I. Borja, A. Aydin, Computational modeling of deformation bands in granular media I: Geological and mathematical framework, *Comput. Meth. Appl. Mech. Engrg.* 193 (2004) 2667–2698.
- [16] R.I. Borja, Computational modeling of deformation bands in granular media, II: numerical simulations, *Comput. Meth. Appl. Mech. Engrg.* 193 (2004) 2699–2718.
- [17] R.I. Borja, K.M. Sama, P.F. Sanz, On the numerical integration of three-invariant elastoplastic constitutive models, *Comput. Meth. Appl. Mech. Engrg.* 192 (2003) 1227–1258.
- [18] R.I. Borja, R.A. Regueiro, Strain localization of frictional materials exhibiting displacement jumps, *Comput. Meth. Appl. Mech. Engrg.* 190 (2001) 2555–2580.
- [19] R.I. Borja, A finite element model for strain localization analysis of strongly discontinuous fields based on standard Galerkin approximations, *Comput. Meth. Appl. Mech. Engrg.* 190 (2000) 1529–1549.
- [20] R.I. Borja, C. Tamagnini, Cam-Clay plasticity Part III: extension of the infinitesimal model to include finite strains, *Comput. Meth. Appl. Mech. Engrg.* 155 (1998) 73–95.
- [21] R.I. Borja, Cam-Clay plasticity, Part II: implicit integration of constitutive equation based on a nonlinear elastic stress predictor, *Comput. Meth. Appl. Mech. Engrg.* 88 (1991) 225–240.
- [22] R.I. Borja, S.R. Lee, Cam-Clay plasticity, Part I: implicit integration of elastoplastic constitutive relations, *Comput. Meth. Appl. Mech. Engrg.* 78 (1990) 49–72.
- [23] E. Burman, Pressure projection stabilizations for Galerkin approximations of Stokes' and Darcy's problem, *Numer. Meth. Part. Differ. Equ.* 24 (2007) 127–143.
- [24] G. Buscarnera, R. Nova, Modelling instabilities in triaxial testing on unsaturated soil specimens, *Int. J. Numer. Anal. Meth. Geomech.* 35 (2011) 179–200.
- [25] C. Callari, A. Abati, Hyperelastic multiphase porous media with strain-dependent retention laws, *Transp. Porous Media* 86 (2011) 155–176.
- [26] C. Callari, F. Armero, A. Abati, Strong discontinuities in partially saturated poroplastic solids, *Comput. Meth. Appl. Mech. Engrg.* 199 (2010) 1513–1535.
- [27] C. Callari, A. Abati, Finite element methods for unsaturated porous solids and their application to dam engineering problems, *Comput. Struct.* 87 (2009) 485–501.
- [28] M. Celia, E. Bouloutas, R. Zarba, A general mass-conservative numerical solution for the unsaturated flow equation, *Water Resour. Res.* 26 (1990) 1483–1496.
- [29] Q. Chen, A. Seifried, J.E. Andrade, J.W. Baker, Characterization of random fields and their impact on the mechanics of geosystems at multiple scales, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 111–248.
- [30] W.F. Chen, D.J. Han, *Plasticity for Structural Engineers*, Springer Verlag, Berlin, Germany, 1988.
- [31] O. Chupin, A.L. Rechenmacher, S. Abedi, Finite strain analysis of nonuniform deformation inside shear bands in sands, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 1651–1666.
- [32] C.J.G. Darnault, J.A. Throop, D.A. DiCarlo, A. Rimmer, T.S. Steenhuis, J.-Y. Parlange, Visualization by light transmission of oil and water contents in transient two-phase flow fields, *J. Cont. Hydr.* 31 (1998) 337–348.
- [33] D.A. DiCarlo, T.W.J. Bauters, T.S. Steenhuis, J.-Y. Parlange, B.R. Bierck, High-speed measurements of three-phase flow using synchrotron X rays, *Water Resour. Res.* 33 (1997) 569–576.
- [34] C.R. Dohrmann, P.B. Bochev, A stabilized finite element method for the Stokes problem based on polynomial pressure projections, *Int. J. Numer. Meth. Fluids* 46 (2004) 183–201.
- [35] M. Dumont, S. Taibi, J.-M. Fleureau, N. Abou-Bekr, A. Saouab, A thermo-hydro-mechanical model for unsaturated soils based on the effective stress concept, *Int. J. Numer. Anal. Meth. Geomech.* 35 (2011) 1299–1317.
- [36] W. Ehlers, T. Graf, M. Ammann, Deformation and localization analysis of partially saturated soil, *Comput. Meth. Appl. Mech. Engrg.* 193 (2004) 2885–2910.
- [37] R.A. Fisher, On the capillary forces in an ideal soil; correction of formulae given by W.B. Haines, *J. Agric. Sci.* 16 (1926) 492–505.
- [38] G. Gallipoli, A. Gens, R. Sharma, J. Vaunat, An elasto-plastic model for unsaturated soil incorporating the effects of suction and degree of saturation on mechanical behaviour, *Géotechnique* 53 (2003) 123–135.
- [39] P. Garnier, R. Angulo-Jaramillo, D.A. DiCarlo, T.W.J. Bauters, C.J.G. Darnault, T.S. Steenhuis, J.-Y. Parlange, P. Baveye, Dual-energy synchrotron X ray measurements of rapid soil density and water content changes in swelling soils during infiltration, *Water Resour. Res.* 34 (1998) 2837–2842.
- [40] G.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1971.
- [41] M.Th. van Genuchten, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.* 44 (1980) 892–898.
- [42] A. Ghassemi, A. Pak, Pore scale study of permeability and tortuosity for flow through particulate media using Lattice Boltzmann method, *Int. J. Numer. Anal. Meth. Geomech.* 35 (2011) 886–901.
- [43] R.J. Glass, T.S. Steenhuis, J.-Y. Parlange, Mechanism for finger persistence in homogeneous, unsaturated, porous media: theory and verification, *Soil Sci.* 148 (1989) 60–70.
- [44] A. Gold, B. Asher, Soil reflectance measurements using a photographic method, *Soil Sci. Am. J.* 40 (1976) 337–341.
- [45] G. Gudehus, Elastoplastische Stoffgleichungen für trockenen Sand, *Ingenieur-Archiv* 42 (1973) 151–169.
- [46] J. Guilleminot, C. Soize, R.G. Ghanem, Stochastic representation for anisotropic permeability tensor random fields, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 1592–1608.
- [47] R. Kohler, G. Hofstetter, A cap model for partially saturated soils, *Int. J. Numer. Anal. Meth. Geomech.* 32 (2007) 981–1004.
- [48] M.G. Jefferies, Nor-Sand: a simple critical state model for sand, *Géotechnique* 43 (1993) 91–103.
- [49] C. Kechavarzi, K. Soga, P. Wiart, Multispectral image analysis method to determine dynamic fluid saturation distribution in two-dimensional three-fluid phase flow laboratory experiments, *J. Cont. Hydr.* 46 (2000) 265–293.
- [50] T.W. Lambe, R.V. Whitman, *Soil Mechanics*, John Wiley & Sons, 1969.
- [51] C. Linder, A. Raina, A strong discontinuity approach on multiple levels to model solids at failure, *Comput. Meth. Appl. Mech. Engrg.* 253 (2013) 558–583.
- [52] C. Linder, D. Rosato, C. Miehe, New finite elements with embedded strong discontinuities for the modeling of failure in electromechanical coupled solids, *Comput. Meth. Appl. Mech. Engrg.* 200 (2011) 141–161.
- [53] C. Linder, F. Armero, Finite elements with embedded strong discontinuities for the modeling of failure in solids, *Int. J. Numer. Meth. Engrg.* 72 (2007) 1391–1433.
- [54] B. Loret, N. Khalili, An effective stress elastic-plastic model for unsaturated porous media, *Mech. Mater.* 34 (2002) 97–116.
- [55] D. Mašin, N. Khalili, A thermo-mechanical model for variably saturated soils based on hypoplasticity, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 1461–1485.
- [56] G. Meschke, S. Grasberger, Numerical modeling of coupled hygro-mechanical degradation of cementitious materials, *J. Engrg. Mech.* 129 (2003) 383–392.
- [57] G. Meschke, Consideration of aging of shotcrete in the context of a 3-D viscoplastic material model, *Int. J. Numer. Meth. Engrg.* 39 (1996) 3123–3143.
- [58] M. Nuth, L. Laloui, Effective stress concept in unsaturated soils: Clarification and validation of a unified framework, *Int. J. Numer. Anal. Meth. Geomech.* 32 (2008) 771–801.
- [59] T.M.H. Le, D. Gallipoli, M. Sanchez, S.J. Wheeler, Stochastic analysis of unsaturated seepage through randomly heterogeneous earth embankments, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 1056–1076.

- [60] A.L. Rechenmacher, S. Abedi, O. Chupin, A.D. Orlando, Characterization of mesoscale instabilities in localized granular shear using digital image correlation, *Acta Geotech.* 6 (2012) 205–215.
- [61] R.A. Regueiro, R.I. Borja, Plane strain finite element analysis of pressure-sensitive plasticity with strong discontinuity, *Int. J. Solids Struct.* 38 (2001) 3647–3672.
- [62] A. Rimmer, D.A. DiCarlo, T.S. Steenhuis, B. Bierck, D. Durnford, J.-Y. Parlange, Rapid fluid content measurement method for fingered flow in an oil-water-sand system using synchrotron X-rays, *J. Cont. Hydr.* 31 (1998) 315–335.
- [63] K.H. Roscoe, J.B. Burland, On the generalized behaviour of 'wet' clay, J. Heyman, F.A. Leckie (Eds.), *Engineering Plasticity*, Cambridge University Press Cambridge, pp. 535–609.
- [64] J.W. Rudnicki, J.R. Rice, Conditions for the localization of deformation in pressure sensitive dilatant materials, *J. Mech. Phys. Solids* 23 (1975) 371–394.
- [65] R.A. Schincariol, F.W. Schwartz, An experimental investigation of variable density flow and mixing in homogeneous and heterogeneous media, *Water Resour. Res.* 26 (1990) 2317–2329.
- [66] R.A. Schincariol, E.E. Herderick, F.W. Schwartz, On the application of image analysis to determine concentration distribution in laboratory experiments, *J. Cont. Hydr.* 12 (1993) 197–215.
- [67] A. Schofield, P. Wroth, *Critical State Soil Mechanics*, McGraw-Hill, New York, 1968.
- [68] R. Suchomel, D. Mašin, Comparison of different probabilistic methods for predicting stability of a slope in spatially variable $c-\phi$ soil, *Comput. Geotech.* 37 (2010) 132–140.
- [69] C. Tamagnini, R. Castellanza, R. Nova, A generalized backward Euler algorithm for the numerical integration of an isotropic hardening elastoplastic model for mechanical and chemical degradation of bonded geomaterials, *Int. J. Numer. Anal. Meth. Geomech.* 26 (2002) 963–1004.
- [70] J. Tejchman, Effect of fluctuation of current void ratio on the shear zone formation in granular bodies within micro-polar hypoplasticity, *Comput. Geotech.* 33 (2006) 29–46.
- [71] V.C. Tidwell, R.J. Glass, X-ray and visible light transmission for laboratory measurements of two-dimensional saturation fields in thin-slab systems, *Water Resour. Res.* 30 (1994) 2873–2882.
- [72] D.M. Tuck, B.R. Bierck, P.R. Jaffe, Synchrotron radiation measurement of multiphase fluid saturation in porous media: experimental technique and error analysis, *J. Cont. Hydr.* 31 (1998) 231–256.
- [73] R. Uzuoka, R.I. Borja, Dynamics of poroelastic solids at finite strain, *Int. J. Numer. Anal. Meth. Geomech.* 36 (2012) 1535–1573.
- [74] P.J. Van Geel, J.F. Sykes, Laboratory and model simulations of a LNAPL spill in a variably-saturated sand: 1. Laboratory experiment and image analysis techniques, *J. Cont. Hydr.* 17 (1994) 1–25.
- [75] J.A. White, R.I. Borja, Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients, *Comput. Meth. Appl. Mechan. Engrg.* 197 (2008) 4353–4366.
- [76] J.A. White, R.I. Borja, Block-preconditioned Newton–Krylov solvers for fully coupled flow and geomechanics, *Comput. Geosci.* 15 (2011) 647–659.
- [77] K.J. Willam, E.P. Warnke, Constitutive model for triaxial behavior of concrete, *Concrete Structures Subjected to Triaxial Stresses*, International Association for Bridges and Structural Engineering, Bergamo, Italy, 1974.
- [78] N. Yoshimoto, R.P. Orense, F. Tanabe, N. Kikkawa, M. Hyodo, Y. Nakata, Measurement of degree of saturation on model ground by digital image processing, *Soils Found.* 51 (2011) 167–177.
- [79] H.W. Zhang, B.A. Schrefler, Gradient-dependent plasticity model and dynamic strain localisation analysis of saturated and partially saturated porous media: one dimensional model, *Eur. J. Mech. A/Solids* 19 (2000) 503–524.
- [80] Y. Zou, A macroscopic model for predicting the relative hydraulic permeability of unsaturated soils, *Acta Geotech.* 7 (2012) 129–137.