Journal of Rock Mechanics and Geotechnical Engineering 17 (2025) 2041–2058

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Hop-to-Hug algorithm: Novel strategy to stable cutting-plane algorithm based on convexification of yield functions

Yanbin Chen [a], Yuanming Lai [b], Enlong Liu [a, c, *]

[a] Institute for Disaster Management and Reconstruction (IDMR), Sichuan University-The Hong Kong Polytechnic University, Chengdu, 610041, China
[b] Northwest Institute of Eco-Environment and Resources, State Key Laboratory of Frozen Soil Engineering, Chinese Academy of Sciences, Lanzhou, 730000, China
[c] State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resources and Hydropower, Sichuan University, Chengdu, 610041, China

ARTICLE INFO

Article history:
Received 9 January 2024
Received in revised form
28 April 2024
Accepted 15 July 2024
Available online 2 August 2024

Keywords:
Elastoplastic model
Constitutive model integration
Cutting-plane algorithm
Geotechnical analysis
Finite element method (FEM)

ABSTRACT

Numerical challenges, incorporating non-uniqueness, non-convexity, undefined gradients, and high curvature, of the positive level sets of yield function $\mathscr{F} > 0$ are encountered in stress integration when utilizing the return-mapping algorithm family. These phenomena are illustrated by an assessment of four typical yield functions: modified spatially mobilized plane criterion, Lade criterion, Bigoni-Piccolroaz criterion, and micromechanics-based upscaled Drucker-Prager criterion. One remedy to these issues, named the "Hop-to-Hug" (H2H) algorithm, is proposed via a convexification enhancement upon the classical cutting-plane algorithm (CPA). The improved robustness of the H2H algorithm is demonstrated through a series of integration tests in one single material point. Furthermore, a constitutive model is implemented with the H2H algorithm into the Abaqus/Standard finite-element platform. Element-level and structure-level analyses are carried out to validate the effectiveness of the H2H algorithm in convergence. All validation analyses manifest that the proposed H2H algorithm can offer enhanced stability over the classical CPA method while maintaining the ease of implementation, in which evaluations of the second-order derivatives of yield function and plastic potential function are circumvented.

1. Introduction

The plasticity theory (Hill, 1998) is one of the most critical branches in solid mechanics to account for the nonlinear responses incorporating the permanent (plastic) deformation and thermodynamic dissipation. Constitutive models built in the framework of plasticity theory often involve complex mathematical descriptions (Hill, 1948; Gurson, 1977; Borja and Lee, 1990; Barlat et al., 2005; Bryant and Sun, 2019; Cornejo et al., 2022), necessitating utilization of the numerical integration methods to facilitate computational simulations. The essence of numerical computation in elastoplastic constitutive models lies in accurately evaluating the incremental plastic strain generated at each faced material point during every sub-increment step. This process involves the mathematical discretization of concepts such as yield surface, plastic flow, and strain hardening/softening within the realm of plasticity theory (Simo and Hughes, 1998; Anandarajah, 2010). The numerical methods for achieving such tasks are referred to as the stress integration (or update) algorithms.

In general, the integration algorithms for plastic constitutive modeling can be categorized into two types: explicit and implicit. Explicit methods are characterized by utilization of the information at the beginning of each integration time step without the need for iteration. Classical explicit schemes include the modified Euler method and the Runge-Kutta family (Sloan, 1987; Sloan et al., 2001; Lloret-Cabot et al., 2021). Due to the ease of implementation, explicit methods are widely used in bounding surface models (Dafalias and Manzari, 2004; Liu et al., 2022) and hypoplastic models (Ding et al., 2015; Jiao et al., 2024). However, explicit algorithms typically suffer from stress drift outside the yield surface, leading to unsatisfactory of the plastic consistency condition.

---

* Corresponding author. State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resources and Hydropower, Sichuan University, Chengdu, 610041, China.
E-mail address: liuenlong@scu.edu.cn (E. Liu).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

https://doi.org/10.1016/j.jrmge.2024.07.009

Consequently, the sub-stepping technique is required to control the progressively accumulated error, and appropriate treatment should be adopted to address the yield surface intersection problem (Sloan et al., 2001). Another widely adopted group is the implicit return-mapping (RM) algorithm family. This family includes the fully implicit closest point projection method (CPPM) (Simo and Ortiz, 1985; Simo and Taylor, 1985; Borja et al., 2001; Tamagnini et al., 2002; Foster et al., 2005; Cardoso and Yoon, 2009; Lin and Li, 2015; Wei et al., 2021; Feng et al., 2024) based on the backward Euler scheme, as well as the semi-implicit cutting-plane algorithm (CPA) (Ortiz and Simo, 1986; Starman et al., 2014; Zhao et al., 2023). The RM algorithms are also referred to as the elastic predictor-plastic corrector methods. An elastic trial step is made at the beginning to check the yield state, and the plastic correction is implemented if the trial state lies outside the admissible domain enclosed by the yield surface. The CPPM utilizes a Newton-Raphson (NR) scheme to solve the fully implicit equations, thereby requiring both the first- and second-order derivatives of the plastic potential function. Accordingly, the implementation of CPPM could be challenging for complex models, but it offers the advantages of high accuracy and unconditional stability. In recent applications (e.g. Lee et al., 2021; Zhou et al., 2022), the line search method has been employed to enhance the convergence of the CPPM. The CPA, on the other hand, implicitly solves the plastic multiplier increments through the NR method and updates the stress and stress-like variables via a forward Euler scheme. This approach circumvents the challenge of dealing with second-order derivatives in the CPPM, facilitating straightforward implementation while satisfying the plastic consistency condition. Consequently, the cutting-plane algorithm has been widely adopted for the numerical implementation of various constitutive models (Areias et al., 2012; Zhu et al., 2016; Ren et al., 2022). For more details on the computational schemes of the elastoplastic model, please refer to the work of Scalet and Auricchio (2018).

It must be acknowledged that although an increasing number of advanced constitutive models are proposed, their numerical implementation on widely accepted computational platforms remains relatively lacking, often preventing these advanced models from penetrating frontline engineering applications. However, in the field of geotechnical engineering, many sophisticated constitutive formulations have been successfully implemented via some plugin interfaces into the commercial finite element platforms like Abaqus (e.g. Zhao et al., 2018; Wei et al., 2021; Zhou et al., 2022), the open-source finite element framework like OpenSees (e.g. Khosravifar et al., 2018; Liu et al., 2018, 2022), and commercial Lagrangian finite-volume program like FLAC3D (e.g. Derbin et al., 2018; Xiao et al., 2024), and among others. The requirement for stability in engineering computations has driven continuous improvements in stress update algorithms to enhance their robustness.

In RM algorithms, the plastic correction utilizes the information provided by the trial state at which the yield function (denoted by $\mathscr{F}$) is temporarily $\mathscr{F} > 0$. It is not difficult to demonstrate that many yield functions, although capable of guaranteeing convexity perfectly when $\mathscr{F} = 0$, often yield non-convexity and even possess mathematically undefined regions in their positive-valued level sets ($\mathscr{F} > 0$) (Coombs and Crouch, 2011; Panteghini and Lagioia, 2018; Golchin et al., 2021). To address these challenges, numerous approaches have been attempted. Panteghini and Lagioia (2018) presented a method to transform general yield functions into quasiconvex ones and applied the convexification technique to a Cam-clay type model with a fully implicit stress update scheme. Golchin et al. (2021) proposed a yield function that is inherently robust, serving as a substitute for the inadequate models as discussed in their paper. A framework called Convex Modular Modeling was established by Suryasentana et al. (2021), which ensures convexity by combining existing convex function modules with the convexity-preserving rules (Boyd and Vandenberghe, 2004).

The present study extends the convexification approach of Panteghini and Lagioia (2018) to a more general form, and proposes an improved semi-implicit stress integration scheme based on the extension, which is named the "Hop-to-Hug" (H2H) algorithm. The novel scheme exhibits enhanced robustness while inheriting the ease of implementation characteristics of the CPA method, making it versatile for application to most existing elastoplastic models. Prior to presenting the H2H algorithm, the paper will first discuss the numerical integration problem of the elastoplastic constitutive model and illustrate the issues with the RM methods through some examples. Then a detailed demonstration of the H2H method with the convexification method is introduced. Finally, a series of numerical validations are conducted on several constitutive models with the H2H algorithm to test its capability.

For clarity, notations and symbols adhere to the following rules. The italic letters (e.g. γ, $\mathscr{F}$), boldface symbols (e.g. $\boldsymbol{\sigma}$, $\boldsymbol{\epsilon}$), and the uppercase black board symbols denote, respectively, the scalars, second-order tensors, and forth-order tensors. Unless otherwise stated, the notations $\{\bullet\}$ and $[\bullet]$ represent vector and matrix, respectively. As for tensor operations, the tensor product is denoted by ($\bullet \otimes \bullet$), while ($\bullet \cdot \bullet$) and ($\bullet : \bullet$) stands for inner and double contraction product, respectively. Some useful notations of Cauchy stress are presented in Appendix A.

2. Problem statement

2.1. Basic equations

Consider an isotropic body $\Omega \subset \Re^{n^{\dim}}$ of elastoplastic continuum, where $1 \leq n^{\dim} \leq 3$ is the space dimension. The body Ω is assumed open and bounded with smooth boundary $\partial\Omega$. The boundary is decomposed into a Dirichlet and a Neumann boundary, denoted respectively $\partial\Omega_{\boldsymbol{u}}$ and $\partial\Omega_{\boldsymbol{\tau}}$, such that

$$\left.\begin{array}{l} \overline{\partial\Omega_{\boldsymbol{u}} \cup \partial\Omega_{\boldsymbol{\tau}}} = \partial\Omega \\ \partial\Omega_{\boldsymbol{u}} \cap \partial\Omega_{\boldsymbol{\tau}} = \varnothing \end{array}\right\} \tag{1}$$

where the overline denotes a closure. Let the time interval of interest be given by $[0, \mathscr{T}] \subset \Re^{+}$, then the displacement field of an arbitrary material point with position $x \in \Omega$ at time $t \in [0, \mathscr{T}]$ can be written as

$$\boldsymbol{u}(\boldsymbol{x}, t) : \overline{\Omega} \times [0, \mathscr{T}] \rightarrow \Re^{n^{\dim}} \tag{2}$$

where

$$\overline{\Omega} = \Omega \cup \partial\Omega$$

The boundary conditions restricted by Eq. (1) are given by

$$\left.\begin{array}{l} \boldsymbol{u} = \overline{\boldsymbol{u}} \ \text{ on } \ \partial\Omega_{\boldsymbol{u}} \times [0, \mathscr{T}] \\ \boldsymbol{\sigma} \cdot \boldsymbol{n} = \overline{\boldsymbol{\tau}} \ \text{ on } \ \partial\Omega_{\boldsymbol{\tau}} \times [0, \mathscr{T}] \end{array}\right\} \tag{3}$$

where $\overline{\boldsymbol{u}}$ and $\overline{\boldsymbol{\tau}}$ are the displacement and surface traction field prescribed to $\partial\Omega_{\boldsymbol{u}}$ and $\partial\Omega_{\boldsymbol{\tau}}$, respectively. The initial conditions are also supplied with

$$\left.\begin{array}{l} \boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}) \ \text{ in } \ \Omega \\ \dot{\boldsymbol{u}}(\boldsymbol{x}, 0) = \boldsymbol{v}_0(\boldsymbol{x}) \ \text{ in } \ \Omega \end{array}\right\} \tag{4}$$

where $\boldsymbol{u}_0$ and $\boldsymbol{v}_0$ are prescribed initial values of displacement and velocity field, respectively. The body deformation is measured by the infinitesimal strain tensor $\boldsymbol{\epsilon}$, defined by the symmetry part of the gradient of $\boldsymbol{u}$:

$$\boldsymbol{\epsilon} = \mathrm{sym}(\nabla\boldsymbol{u}) \tag{5}$$

The internal stress response caused by the prescribed boundary conditions and body force fields is measured by the Cauchy stress tensor $\boldsymbol{\sigma}$. Thus, the local form of conservation of momentum is stated as

$$\nabla\cdot\boldsymbol{\sigma} + \rho\boldsymbol{b} = \rho\ddot{\boldsymbol{u}} \quad \text{in} \quad \Omega\times[0,\mathscr{T}] \tag{6}$$

where ρ is the mass density and $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{x}, t)$ stands for the body force per unit of mass. The term $\rho\ddot{\boldsymbol{u}}$ vanishes when quasi-static loading is considered.

The governing equations in global (weak) form are given by the principle of virtual work expressed in the current configuration:

$$\int_{\Omega}(\boldsymbol{\sigma} : \delta\boldsymbol{\epsilon} + \rho\,\delta\boldsymbol{u}\cdot\ddot{\boldsymbol{u}})\mathrm{d}\Omega = \int_{\Omega}\rho\,\delta\boldsymbol{u}\cdot\boldsymbol{b}\,\mathrm{d}\Omega + \int_{\partial\Omega}\delta\boldsymbol{u}\cdot\overline{\boldsymbol{\tau}}\,\mathrm{d}(\partial\Omega) \tag{7}$$

where $\delta\boldsymbol{u}$ and $\delta\boldsymbol{\epsilon} = \mathrm{sym}\,(\nabla\delta\boldsymbol{u})$ are the virtual displacement and strain filed, respectively.

If a set of constitutive equations relate stress $\boldsymbol{\sigma}$ to strain $\boldsymbol{\epsilon}$ are meticulously devised, the partial differential equations Eq. (6), alongside the initial conditions Eq. (4) and boundary conditions Eq. (3), yields an initial boundary value problem (IBVP) for the displacement field $\boldsymbol{u}$ (Simo and Hughes, 1998; Anandarajah, 2010; de Borst et al., 2012). To solve the IBVPs, the weak form Eq. (7) could be discretized by a variety of numerical methods, e.g. finite element method (FEM), material point method (MPM), and smoothed particle hydrodynamics (SPH).

2.2. Elastoplastic constitutive equations

It is assumed that the strain tensor $\delta\boldsymbol{\epsilon}$ conforms to an additive decomposition such that we have

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\mathrm{e}} + \boldsymbol{\epsilon}^{\mathrm{p}} \tag{8}$$

where $\boldsymbol{\epsilon}^{\mathrm{e}}$ and $\boldsymbol{\epsilon}^{\mathrm{p}}$ are the elastic (reversible) strain and plastic (irreversible) strain components, respectively.

The elastic response is described by the generalized Hooke's law as follows:

$$\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\epsilon}^{\mathrm{e}} = \mathbb{C} : (\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^{\mathrm{p}}) \tag{9}$$

Classical rate-independent plasticity recognizes the plastic flow as an irreversible process, which depends on the history of the total strain $\boldsymbol{\epsilon}$, the plastic strain $\boldsymbol{\epsilon}^{\mathrm{p}}$, and a set of strain-like internal variables $\boldsymbol{\xi}$. The variables $\boldsymbol{\xi}$ are often known as hardening parameters, which can be scalars and/or higher-order tensors. Without loss of generality, the hardening parameters $\boldsymbol{\xi}$ are defined as a vector of scalar variables in the present study. The stress $\boldsymbol{\sigma}$ and a set of stress-like internal variables $\boldsymbol{\chi}$ thermodynamically conjugate to $\boldsymbol{\epsilon}^{\mathrm{p}}$ and $\boldsymbol{\xi}$, respectively. The admissible region of the state $(\boldsymbol{\sigma}, \boldsymbol{\chi})$ is defined by a convex set as

$$E_{\boldsymbol{\sigma}} = \{(\boldsymbol{\sigma}, \boldsymbol{\chi}) : \mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \le 0\} \tag{10}$$

where $\mathscr{F} = \mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi})$ is a convex scalar function, which is termed as the yield function. The interior of $E_{\boldsymbol{\sigma}}$ denoted by

$$\mathrm{int}(E_{\boldsymbol{\sigma}}) = \{(\boldsymbol{\sigma}, \boldsymbol{\chi}) : \mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) < 0\} \tag{11}$$

Eq. (11) is called the elastic region, inside which the state $(\boldsymbol{\sigma}, \boldsymbol{\chi})$ is purely elastic. It is identified as a plastic behavior when the state $(\boldsymbol{\sigma}, \boldsymbol{\chi})$ lies on the boundary of $E_{\boldsymbol{\sigma}}$, which is indicated by

$$\partial E_{\boldsymbol{\sigma}} = \{(\boldsymbol{\sigma}, \boldsymbol{\chi}) : \mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) = 0\} \tag{12}$$

The set $\partial E_{\boldsymbol{\sigma}}$ defined by $\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) = 0$ is also termed the yield surface. It is worth noting that the components of $\boldsymbol{\chi}$ are typically material quantities, and their evolution will be constrained within reasonable bounds, with no impact on the convexity of the elastic regime. In theory, the complement of set $E_{\boldsymbol{\sigma}}$ is non-admissible. However, it shall be expounded on in later sections that the trial stress state and the intermediate state in the iteration process often temporarily violate this theoretical definition whilst the elastic predictor-plastic corrector numerical stress integration strategy is adopted.

The plastic flow rule and hardening law are assumed as

$$\left.\begin{aligned} \dot{\boldsymbol{\epsilon}}^{\mathrm{p}} &= \dot{\gamma}\frac{\partial\mathscr{G}}{\partial\boldsymbol{\sigma}} = \dot{\gamma}\boldsymbol{r}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \\ \dot{\boldsymbol{\chi}} &= -\dot{\gamma}\boldsymbol{h}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \end{aligned}\right\} \tag{13}$$

where $\mathscr{G} = \mathscr{G}(\boldsymbol{\sigma}, \boldsymbol{\chi})$ is the scalar-valued convex plastic potential, and $\boldsymbol{r}$ and $\boldsymbol{h}$ indicate the plastic flow direction and generalized hardening moduli, respectively. It is noted that non-smooth potential functions can also be treated with the smoothing approximation or the rule of Koiter's generalization (de Borst et al., 2012; Karaoulanis, 2013). The algorithms introduced in the following can be readily extended to non-smooth models by these techniques. Therefore, this work will primarily focus on the smooth case. The plasticity is referred to as associative if the plastic potential $\mathscr{G}$ aligns with the yield criterion $\mathscr{F}$. Otherwise, it is known as the non-associative plasticity. The non-negative scalar function $\dot{\gamma}$ is the plastic consistency parameter, which satisfies the Karush-Kuhn-Tucker (KKT) complementarity conditions:

$$\left.\begin{aligned} &\dot{\gamma} \ge 0 \\ &\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \le 0 \\ &\dot{\gamma}\mathscr{F} = 0 \end{aligned}\right\} \tag{14}$$

and the consistency requirement:

$$\dot{\gamma}\dot{\mathscr{F}} = 0 \tag{15}$$

For further interpretation of the KKT and consistency conditions, readers can find the work of Simo and Hughes (1998).

2.3. Numerical stress integration

Numerical analysis of structures formed by nonlinear materials often requires an implicit iterative scheme at the global (structure) level through Newton's method (or its variants). The goal of forming the global stiffness matrix then necessitates the stress integration at the local level (e.g. Gauss point in the finite element method) to evaluate the local consistent tangent matrix.

Remark 1. The additive operator split method can be introduced for elastoplastic problems:

$$\left.\begin{array}{ccccc} \text{Total} & \equiv & \text{Elastic predictor} & + & \text{Plastic corrector} \\ \dot{\boldsymbol{\epsilon}} = \mathrm{sym}(\nabla\dot{\boldsymbol{u}}) & & \dot{\boldsymbol{\epsilon}} = \mathrm{sym}(\nabla\dot{\boldsymbol{u}}) & & \dot{\boldsymbol{\epsilon}} = 0 \\ \dot{\boldsymbol{\epsilon}}^{\mathrm{p}} = \dot{\gamma}\boldsymbol{r}(\boldsymbol{\sigma}, \boldsymbol{\chi}) & & \dot{\boldsymbol{\epsilon}}^{\mathrm{p}} = 0 & & \dot{\boldsymbol{\epsilon}}^{\mathrm{p}} = \dot{\gamma}\boldsymbol{r}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \\ \dot{\boldsymbol{\chi}} = -\dot{\gamma}\boldsymbol{h}(\boldsymbol{\sigma}, \boldsymbol{\chi}) & & \dot{\boldsymbol{\chi}} = 0 & & \dot{\boldsymbol{\chi}} = -\dot{\gamma}\boldsymbol{h}(\boldsymbol{\sigma}, \boldsymbol{\chi}) \end{array}\right\} \tag{16}$$

According to the Lie-Trotter product formula (Hall, 2015), the splitting can lead to the elastic predictor-plastic corrector methodology. The elastic predictor "freezes" the plastic flow and hardening evolution and presumes that the strain rate is purely elastic. Subsequently, the plastic corrector utilizes the elastic predictor as the initial condition and maps the inadmissible stress state back to the updated yield surface. Consider a local material point at the converged global computational step n. The state of the point is characterized by $\boldsymbol{\epsilon}_n$, $\boldsymbol{\epsilon}_n^{\mathrm{p}}$, $\boldsymbol{\sigma}_n$, and $\boldsymbol{\chi}_n$, which satisfies $\mathscr{F}(\boldsymbol{\sigma}_n, \boldsymbol{\chi}_n) \le 0$ as well. The commencement of the integration is a prescribed incremental strain $\Delta\boldsymbol{\epsilon} = \mathrm{sym}\,(\nabla(\Delta\boldsymbol{u}))$ driving the calculation for the solution of state in the step $n+1$. The elastic predictor, which is also termed as trial stress state, is obtained by elastically integrating over $\Delta\boldsymbol{\epsilon}$:

$$\boldsymbol{\sigma}^{\mathrm{tri}} = \mathbb{C} : \left(\boldsymbol{\epsilon}_{n+1} - \boldsymbol{\epsilon}_n^{\mathrm{p}}\right) \tag{17}$$

where the superscript "tri" stands for trial state and $\boldsymbol{\epsilon}_{n+1} = \boldsymbol{\epsilon}_n + \Delta\boldsymbol{\epsilon}$. The trial stress $\boldsymbol{\sigma}^{\mathrm{tri}}$ and the "frozen" internal variables $\boldsymbol{\chi}_n$ are used to check whether the current step is elastic or inelastic through

$$\left.\begin{aligned} &\mathscr{F}\left(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n\right) \le 0 \implies \text{Elastic step} \\ &\mathscr{F}\left(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n\right) > 0 \implies \text{Elastoplastic step} \end{aligned}\right\} \tag{18}$$

If the current step is confirmed elastic, then the trial state is accepted for step $n+1$:

$$\left.\begin{aligned} \boldsymbol{\sigma}_{n+1} &= \boldsymbol{\sigma}^{\mathrm{tri}} \\ \boldsymbol{\epsilon}_{n+1}^{\mathrm{p}} &= \boldsymbol{\epsilon}_n^{\mathrm{p}} \\ \boldsymbol{\chi}_{n+1} &= \boldsymbol{\chi}_n \end{aligned}\right\} \tag{19}$$

Yet for the elastoplastic step, the state $(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n)$ lies outside the admissible region and violates the KKT and consistency conditions. Consequently, the plastic corrector problem in incremental form is raised as follows:

$$\left.\begin{aligned} &\boldsymbol{\sigma}_{n+1} = \boldsymbol{\epsilon}_{n+1} - \boldsymbol{\epsilon}_{n+1}^{\mathrm{p}} \\ &\boldsymbol{\epsilon}_{n+1}^{\mathrm{p}} = \boldsymbol{\epsilon}_n^{\mathrm{p}} + \Delta\gamma\boldsymbol{r}(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1}) \\ &\boldsymbol{\chi}_{n+1} = \boldsymbol{\chi}_n - \Delta\gamma\boldsymbol{h}(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1}) \\ &\mathscr{F}(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1}) = 0 \end{aligned}\right\} \tag{20}$$

Eq. (20) with unknowns $\boldsymbol{\sigma}_{n+1}$, $\boldsymbol{\epsilon}_{n+1}^{\mathrm{p}}$, $\boldsymbol{\chi}_{n+1}$, and $\Delta\gamma$ is solved by numerical iterative method, which makes the stress state "return" to the yield surface and restores the required conditions. As discussed previously, CPPM and CPA are two commonly used algorithms for solving plastic correction problems, due to their generally good performance (Borja and Lee, 1990; Zhu et al., 2016). Both techniques rely on the NR method for iterative solution; consequently, the initial state (i.e. the trial state), acting as the starting point for the NR iterations, will directly influence the proper operation of the stress update program. Specifically, the existence and convexity of the surfaces corresponding to Eq. (21) play a critical role in determining the convergence of the RM algorithms.

$$\left.\begin{aligned} \mathscr{F}_{n+1}^{(i)} &= \mathscr{F}\left(\boldsymbol{\sigma}_{n+1}^{(i)}, \boldsymbol{\chi}_{n+1}^{(i)}\right) > 0 \\ \mathscr{G}_{n+1}^{(i)} &= \mathscr{G}\left(\boldsymbol{\sigma}_{n+1}^{(i)}, \boldsymbol{\chi}_{n+1}^{(i)}\right) > 0 \end{aligned}\right\} \tag{21}$$

where the superscript (i) denotes the state corresponding to the i th iteration in the plastic corrector procedure, and $i = 0$ corresponds to the trial stress state.

Remark 2. In the spirit of CPA, the plastic corrector problem is rephrased as a system of ODEs:

$$\left.\begin{aligned} \frac{\mathrm{d}\boldsymbol{\sigma}(\Delta\gamma)}{\mathrm{d}\Delta\gamma} &= -\mathbb{C}(\Delta\gamma) : \boldsymbol{r}(\boldsymbol{\sigma}(\Delta\gamma), \boldsymbol{\chi}(\Delta\gamma)\,) \\ \frac{\mathrm{d}\boldsymbol{\chi}(\Delta\gamma)}{\mathrm{d}\Delta\gamma} &= -\boldsymbol{h}(\boldsymbol{\sigma}(\Delta\gamma), \boldsymbol{\chi}(\Delta\gamma)\,) \end{aligned}\right\} \text{ s.t. } \left\{\boldsymbol{\sigma}(\Delta\gamma), \boldsymbol{\chi}(\Delta\gamma)\,\right\}\Big|_{\Delta\gamma=0} = \left\{\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n\right\} \tag{22}$$

where

$$\boldsymbol{\sigma}(\Delta\gamma) = \mathbb{C}(\Delta\gamma) : \left[\boldsymbol{\epsilon}_{n+1} - \boldsymbol{\epsilon}_{n+1}^{\mathrm{p}}(\Delta\gamma)\right]$$

The solution is a function with $\mathfrak{D} \subset \mathfrak{R}^+$:

$$Y : \Delta\gamma \in \mathfrak{D} \to \{\boldsymbol{\sigma}(\Delta\gamma), \boldsymbol{\chi}(\Delta\gamma)\} \tag{23}$$

Then, the KKT and consistency conditions are enforced by finding the root $\Delta\gamma^* \in \mathfrak{D}$, such that we have

$$\mathscr{F}^* = \mathscr{F}[\boldsymbol{\sigma}(\Delta\gamma^*), \boldsymbol{\chi}(\Delta\gamma^*)] = 0 \tag{24}$$

By performing the Taylor expansion to $\mathscr{F}$ around the state of i th iteration, one arrives at

$$\mathscr{F}^* = \mathscr{F}^{(i)} + (\nabla_{\Delta\gamma}\mathscr{F})^{(i)}\left[\Delta\gamma^* - \Delta\,\gamma^{(i)}\right] + \mathscr{O}\left[(\Delta\gamma)^2\right] \tag{25}$$

where the gradient term takes advantage of the linearized segment of Y and is given by

$$\nabla_{\Delta\gamma}\mathscr{F} \equiv \frac{\mathrm{d}\mathscr{F}}{\mathrm{d}\Delta\gamma} = -\left(\frac{\partial\mathscr{F}}{\partial\boldsymbol{\sigma}}\right) : \mathbb{C} : \boldsymbol{r} - \left(\frac{\partial\mathscr{F}}{\partial\boldsymbol{\chi}}\right) \boldsymbol{\cdot} \boldsymbol{h} \tag{26}$$

Dropping terms of order higher than linear in the right-hand side of Eq. (25), it leads to a newly approximated value:

$$\Delta\gamma^{(i+1)} = \Delta\gamma^{(i)} - \frac{\mathscr{F}^{(i)}}{(\nabla_{\Delta\gamma}\mathscr{F})^{(i)}} \tag{27}$$

The iterative Eq. (27) is in the form of the acclaimed NR method. Thereby, Eq. (27) can expeditiously furnish an adequately precise result, provided that the following conditions are at least satisfied:

1. $(\nabla_{\Delta\gamma}\mathscr{F})^{-1}$ exists and is bounded $\left|(\nabla_{\Delta\gamma}\mathscr{F})^{-1}\right| \le \beta < \infty$ for $\Delta\gamma \in \mathfrak{D}$, where β is a real number; and
2. The first guess $\Delta\gamma^{(0)} = 0$ is sufficiently close to the solution $\Delta\gamma^*$.

As pointed out above, the yield surface $\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) = 0$ is generally

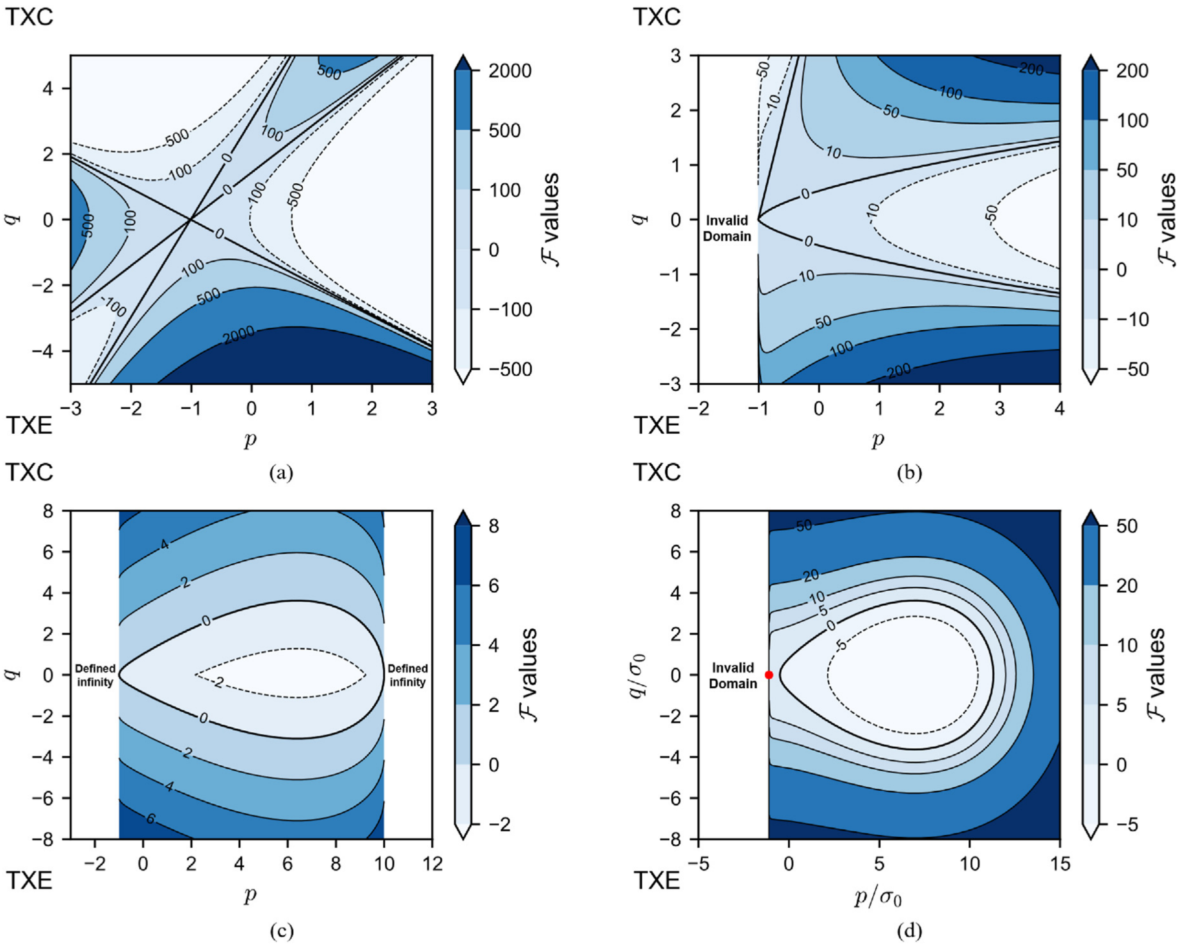


Fig. 1. Level sets of (a) MSMP yield function with $\phi_c = 30^\circ$ and $a^S = 10$; (b) Lade yield function with $\eta_1 = 30$, $m^L = 0.75$, and $a^L = 10$; (c) BP yield function with $M = 0.5$, $p_c = 10$ MPa, $c = 1$ MPa, $\alpha^{BP} = 0.2$, $m^{BP} = 2.5$, $\beta^{BP} = 1.5$, and $\gamma^{BP} = 0.75$; (d) UDP yield function with $\alpha = 0.3$, $\phi = 0.2$. Notes: The red point in (d) indicates a singular point. The abbreviations "TXC" and "TXE" denote respectively the triaxial-compression and triaxial-extension side.

convex, not to mention its existence in mathematics. However, it is usually not the case for the non-zero level set $\mathcal{F} > 0$. To provide a more intuitive illustration of the matter at hand, the positive-valued level sets of some yield functions are plotted in the meridian plane as shown in Fig. 1. These investigated yield functions include the modified spatially mobilized plane (MSMP) criterion (Matsuoka, 1974; Matsuoka and Nakai, 1985), Lade criterion (Lade, 1977), Bigoni-Piccolroaz (BP) criterion (Bigoni and Piccolroaz,

Table 1
Mathematical formulation and model parameters of the analyzed yield functions.

Yield function	Mathematical formulation	Model parameters
MSMP	$\mathcal{F} = \bar{I}_1\bar{I}_2 - k^S\bar{I}_3, \bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - a^S p_{\mathrm{atm}}\boldsymbol{\delta}$	a^S reflects the cohesion; $k^S = \dfrac{9-\sin^2\phi_c}{1-\sin^2\phi_c}$, ϕ_c is the Mohr-Coulomb friction angle
Lade	$\mathcal{F} = \bar{I}_1^3 - \bar{I}_3\left[27 + \eta_1\left(\dfrac{\bar{I}_1}{p_{\mathrm{atm}}}\right)^{-m^L}\right], \bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - a^L p_{\mathrm{atm}}\boldsymbol{\delta}$	a^L reflects the cohesion; η_1 and m^L control the nonlinear pressure dependence
BP	$\mathcal{F} = f(p) + \dfrac{q}{g(\theta)}f(p) =$ $\begin{cases} -Mp_c\sqrt{(\Phi - \Phi^{m^{BP}})(2(1-\alpha^{BP})\Phi + \alpha^{BP})} \text{ (if } \Phi \in [0,1]) \\ +\infty (\text{if } \Phi \notin [0,1]) \end{cases}$ $g(\theta) = 1/\cos\left[\beta^{BP}\dfrac{\pi}{6} - \dfrac{1}{3}\arccos(\gamma^{BP}\cos 3\theta)\right]$ $\Phi = \dfrac{p+c}{p_c+c}$	$\underbrace{M>0, p_c>0, c\geq 0, 0<\alpha^{BP}<2, m^{BP}>1,}_{\text{defining } f(p)}\ \underbrace{0\leq\beta^{BP}\leq 2, 0\leq\gamma^{BP}<1}_{\text{defining } g(\theta)}$ M reflects the pressure dependence; p_c and c are the yield strengths under isotropic compression and tension, respectively; α^{BP}, β^{BP}, m^{BP}, and γ^{BP} control the shape of yield surface
UDP	$\mathcal{F} = \dfrac{q^2}{\sigma_0^2} - \left(1 + 3\alpha\dfrac{p}{\sigma_0}\right)^2 + 2\phi\left(1 + 3\alpha\dfrac{p}{\sigma_0}\right)\cosh\dfrac{1}{2\alpha}\ln\left(1 + 3\alpha\dfrac{p}{\sigma_0}\right) - \phi^2$	α, σ_0 are the material parameters of the solid media, ϕ is the porosity

Note: The MSMP and Lade yield functions have been rewritten from their original expressions to steer clear of the singularities arising from the fractional form. Herein, $\bar{I}_1$, $\bar{I}_2$, and $\bar{I}_3$ are the stress invariants evaluated from the shifted stress $\bar{\boldsymbol{\sigma}}$. The second-order tensor $\boldsymbol{\delta}$ is the Kronecker delta. The atmospheric pressure $p_{\mathrm{atm}} = -101.3$ kPa is adopted as a reference pressure. The definition of the stress invariants is given in Appendix A. Please refer to as the literature for a detailed understanding of the significance of the model parameters mentioned above.

2004), and a micromechanics-based upscaled Drucker-Prager (UDP) criterion (Durban et al., 2010). The mathematical expressions and their associated parameters are summarized in Table 1.

By inspection of the level set contours in Fig. 1a and b, the non-uniqueness and non-convexity of both the MSMP and Lade yield functions can be recognized. The zero-level set $\mathscr{F} = 0$ contains more than one branch, which also divides the elastic domain (where $\mathscr{F} < 0$) into several parts. However, only the branches that embrace the elastic domain in which $p > 0$ should be regarded as the real yield surface. Unfortunately, the negative gradients of the yield function in the regions $\mathscr{F} > 0$ could direct to the undesired yield surfaces, which could lead the return-mapping step towards these undesired directions. As shown in Fig. 1c, though the level sets of the BP yield function avoid being non-unique and non-convex, the gradients are undefined where $p > p_{\mathrm{c}}$ and $p < c$ since the $\mathscr{F}$ value is defined infinite in these domains. For the UDP yield function illustrated in Fig. 1d, the existence of a logarithmic function in the mathematical expression gives rise to a semi-infinite invalid domain and a singular point. Apart from that, the curvature of the contours becomes increasingly pronounced as they approach the invalid region.

Upon analyzing the level sets of yield functions presented above and considering the convergence conditions mentioned in Remark 2, it is possible to deduce a multitude of probable complications that might arise concerning the convergence of the CPA method as follows:

(1) Non-uniqueness. Multiple branches of yield surfaces lead to multiple elastic domains, which may cause the return-mapping corrected stress state located at the undesired surface branch.
(2) Non-convexity. Negative gradients at the trial state and the intermediate states may not well direct the exact yield surface. This could slow down the convergence rate (due to more iterations) and may even lead the updated state to the undesired yield surface.
(3) Gradient undefined. The trial state or the intermediate states positioned in the domain where gradients are undefined will immediately lead to an utter collapse of the CPA.
(4) High curvature. The return-mapping step in the high-curvature region may slow down the convergence (Zhao et al., 2023).

In computational procedures in practice, challenges are often circumvented by either diminishing the computation step size or leveraging sub-stepping techniques. A sufficiently small step size may effectively confine the trial stress state within a "safe" region where the convergence conditions of the CPA method hold. However, such a scheme would evidently increase the computational costs. Nevertheless, the small step size is not a cure-all for all numerical issues. For instance, if a loading path of hydrostatic compression starts from $p = 10$ and $q = 0$ in Fig. 1c is considered. This path will invariably culminate in a collapse of the plastic corrector procedure. Additionally, Panteghini and Lagioia (2018) identified another common phenomenon in the IBVPs. Specifically, in models such as the Cam-clay model (Roscoe and Burland, 1968; Schofield and Wroth, 1968), the initial yield surface's magnitude is regulated by the consolidation stress, which in turn is determined by the geostatic stress. Near the surface of the foundation, the consolidation stress in the soil is relatively low, resulting in a smaller initial yield surface. In such cases, even minuscule increments of applied load can lead to a trial stress state that is far away from the initial elastic domain and therefore may not satisfy the convergence conditions of the CPA method.

In order to address the diverse boundary conditions encountered in numerical computation, it is crucial for computer programs to possess a high level of robustness and reliability, enabling them to effectively handle numerical challenges. Therefore, the objective of this study is to introduce a robust and efficient numerical integration algorithm, which can be considered as an enhanced version of the CPA method. The proposed methodology, built upon the convexification of the yield function, possesses the aptitude to tackle the above-mentioned convergence problems. In addition, it boasts a straightforward implementation, facilitating its seamless integration into existing numerical simulation methods.

3. Hop to hug: A robust cutting-plane algorithm

This section proposes a robust cutting-plane algorithm H2H to tackle irritating convergence issues based on the so-called convexification technique. The consistent tangent operator for H2H is proposed as well.

3.1. Convexification technique

The convexification technique is a method to construct a quasi-convex yield function f against the original non-quasi-convex yield function $\mathscr{F}$ (Panteghini and Lagioia, 2018). First of all, a reference point, denoted by $\boldsymbol{\sigma}^{\mathrm{ref}}$, is selected on the main diagonal of the Haigh–Westergaard (HW) stress space (Hill, 1998), which must always remain within the elastic region bounded by the yield surface. The original yield surface $\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) = 0$ is then subjected to a homothetic transformation where the reference point $\boldsymbol{\sigma}^{\mathrm{ref}}$ is the homothetic center and $z \in \Re$ is the scaling ratio (Coxeter, 1969; Meserve, 2014). The transformation is expressed as

$$\left.\begin{aligned} \tilde{\boldsymbol{\sigma}} &= \boldsymbol{\sigma}^{\mathrm{ref}} + z\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\mathrm{ref}}\right) \\ \boldsymbol{\sigma} &= \boldsymbol{\sigma}^{\mathrm{ref}} + \frac{1}{z}\left(\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^{\mathrm{ref}}\right) \end{aligned}\right\} \tag{28}$$

where the tilde symbol $\tilde{\bullet}$ indicates the quantity is transformed from the original yield surface. It is also noted that $\boldsymbol{\sigma}^{\mathrm{ref}} = \varsigma^{\mathrm{ref}}\boldsymbol{\delta}$ where $\varsigma^{\mathrm{ref}} = \mathrm{tr}(\boldsymbol{\sigma}^{\mathrm{ref}}/3)$ and $\boldsymbol{\delta}$ is the Kronecker delta. Thereby, a group of homothetic surfaces is obtained by substituting the second equation in Eq. (28) into the original yield surface:

$$\tilde{\mathscr{F}} = \tilde{\mathscr{F}}(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}, z) = 0 \tag{29}$$

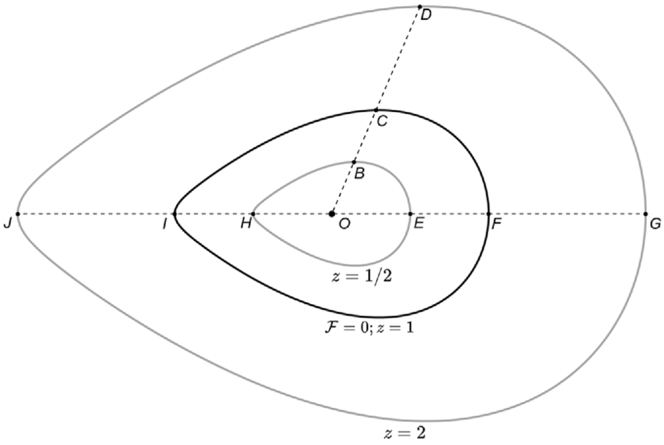


Fig. 2. Schematic diagram of the homothetic transformation. The reference point O is picked as the homothetic center. The grey pictures are transformed from the surface $\mathscr{F} = 0$ with scaling ratios $z = 1/2$ and $z = 2$, respectively. The transformation maintains the shape of the original surface, such that $OB/OC = OE/OF = OH/OI = 1/2$ and $OD/OC = OG/OF = OJ/OI = 2$.

A schematic diagram of geometry is illustrated in Fig. 2. Several crucial insights can be gleaned from this homothety. For instance, the homothety preserves the shape of the original yield surface. Moreover, the homothety with the scaling ratio $z = 1$ is the identity transformation, while the one with a $z < 0$ (or $z > 0$) indicates a surface inside (or outside) the original yield surface. More importantly, when given a state $\boldsymbol{\chi}$ that determines the shape of surface $\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}) = 0$ and a chosen reference point, every specific $z \in \Re$ and $z \neq 0$ determines a transformed surface. This indicates that the implicit function $z = z(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi})$ represents the homothety. Hence, it is sufficient to deduce that a new yield function f can be defined by taking advantage of the implicit function $z = z(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi})$:

$$f = f(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) \equiv z(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) - 1 \tag{30}$$

The newly proposed yield function f is quasi-convex (Bigoni and Piccolroaz, 2004). This is substantiated by the observation that each of its level sets, $L_a = (\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) : f(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) \leq a$, maintains convexity for $a \in \Re$ and $a \neq -1$. Furthermore, the yield function, f, delineates surfaces that bear resemblance to the original surface. In particular, $f = 0$ coincides with the original yield surface $\mathscr{F} = 0$. And f resides either within or beyond the original yield surface, contingent on the $(f < 0)/(f > 0)$ value. It is noteworthy that deriving an explicit expression for the yield function f is not invariably feasible, particularly for the original yield functions of intricate forms. Nevertheless, in practical applications, it proves adequate to resort to numerical methods for determining the value of function f, even in the absence of its explicit expression. The first-order derivatives of f can be evaluated from Eq. (29):

$$\left.\begin{aligned}
\frac{\partial f}{\partial \tilde{\boldsymbol{\sigma}}} &= -\frac{\dfrac{\partial \tilde{\mathscr{F}}}{\partial \tilde{\boldsymbol{\sigma}}}}{\dfrac{\partial \tilde{\mathscr{F}}}{\partial z}} = \frac{z\dfrac{\partial \tilde{\mathscr{F}}}{\partial \boldsymbol{\sigma}}}{\dfrac{\partial \tilde{\mathscr{F}}}{\partial \boldsymbol{\sigma}} : \left(\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^{\mathrm{ref}}\right)} \\
\frac{\partial f}{\partial \boldsymbol{\chi}} &= -\frac{\dfrac{\partial \tilde{\mathscr{F}}}{\partial \boldsymbol{\chi}}}{\dfrac{\partial \tilde{\mathscr{F}}}{\partial z}} = \frac{z^2\dfrac{\partial \tilde{\mathscr{F}}}{\partial \boldsymbol{\chi}}}{\dfrac{\partial \tilde{\mathscr{F}}}{\partial \boldsymbol{\sigma}} : \left(\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^{\mathrm{ref}}\right)}
\end{aligned}\right\} \tag{31}$$

The methodology of determining a new quasi-convex yield function via a homothety of the original yield surface is referred to as the convexification technique. The convexification technique can be applied in practical scenarios by the transformation in principal stress space:

$$\tilde{\boldsymbol{\sigma}}_A = \varsigma^{\mathrm{ref}} + z\left(\boldsymbol{\sigma}_A - \varsigma^{\mathrm{ref}}\right) \quad \forall A = 1, 2, 3 \tag{32}$$

Owing to its conciseness and versatility, the convexification technique is also adept at handling plastic potential functions. For a non-associated elastoplastic model, the convexification technique should be applied concurrently to both the yield function and the plastic potential function. Following the convexification technique above, it requires minimal effort to obtain a homothetic surface group $\tilde{\mathscr{G}}(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}, z')$ and a quasi-convex plastic potential $g(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) = z'(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi}) - 1$ from the original plastic potential $\mathscr{G}(\boldsymbol{\sigma}, \boldsymbol{\chi})$, given a reference point and the scaling factor z'. A primary interest lies in the first-order derivatives of $g(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\chi})$ by

$$\left.\begin{aligned}
\frac{\partial g}{\partial \tilde{\boldsymbol{\sigma}}} &= -\frac{\dfrac{\partial \tilde{\mathscr{G}}}{\partial \tilde{\boldsymbol{\sigma}}}}{\dfrac{\partial \tilde{\mathscr{G}}}{\partial z}} = \frac{z'\dfrac{\partial \tilde{\mathscr{G}}}{\partial \boldsymbol{\sigma}}}{\dfrac{\partial \tilde{\mathscr{G}}}{\partial \boldsymbol{\sigma}} : \left(\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^{\mathrm{ref}}\right)} \\
\frac{\partial g}{\partial \boldsymbol{\chi}} &= -\frac{\dfrac{\partial \tilde{\mathscr{G}}}{\partial \boldsymbol{\chi}}}{\dfrac{\partial \tilde{\mathscr{G}}}{\partial z}} = \frac{z'^2\dfrac{\partial \tilde{\mathscr{G}}}{\partial \boldsymbol{\chi}}}{\dfrac{\partial \tilde{\mathscr{G}}}{\partial \boldsymbol{\sigma}} : \left(\tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^{\mathrm{ref}}\right)}
\end{aligned}\right\} \tag{33}$$

The quasi-convex functions then serve as a stepping stone to overcome the convergence issue of numerical integration mentioned in Section 2.3. Based on the quasi-convex yield function, the improved CPA named the H2H algorithm is proposed below.

3.2. Hop-to-Hug algorithm

Let us return to the plastic corrector problem in Section 2.3, where a trial stress state $(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n)$ that falls outside the yield surface $\mathscr{F}(\boldsymbol{\sigma}, \boldsymbol{\chi}_n)$ needs to be corrected to obtain a convergent result $(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1})$ that satisfies $\mathscr{F}(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1}) = 0$. By virtue of the homothety Eq. (29), it is always possible to obtain a surface passing through an arbitrary point (excluding the reference point) in the stress space, along with its corresponding z and f values. Therefore, the plastic corrector problem can be reformulated as starting from the trial stress state located on $\tilde{\mathscr{F}}(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n, f = f^{\mathrm{tri}} > 0) = 0$, to seek a convergent solution $(\boldsymbol{\sigma}_{n+1}, \boldsymbol{\chi}_{n+1})$ that satisfies $\tilde{\mathscr{F}}(\boldsymbol{\sigma}^{\mathrm{tri}}, \boldsymbol{\chi}_n, f = 0) = 0$.

According to Remark 2, a mapping Υ as expressed in Eq. (23) exists for a plastic corrector problem. As such, the desired solution can be acquired by finding the root $\Delta\gamma^* \in \mathfrak{D} \subset \Re^+$:

$$f^* = f[\boldsymbol{\sigma}(\Delta\gamma^*), \boldsymbol{\chi}(\Delta\gamma^*)] = 0 \tag{34}$$

By employing operations analogous to those in Eqs. (25)–(27), an iterative scheme based on the quasi-convex yield function f can be constructed:

$$\Delta\gamma^{(i+1)} = \Delta\gamma^{(i)} - \frac{f^{(i)}}{\left(\nabla_{\Delta\gamma} f\right)^{(i)}} \tag{35}$$

where the derivation of the gradient term $\nabla_{\Delta\gamma} f$ also avails itself of the linearized segment of Eq. (23). The gradient $\nabla_{\Delta\gamma} f$ is given by

$$\nabla_{\Delta\gamma} f = -\frac{\partial f}{\partial \boldsymbol{\sigma}} : \mathbb{C} : \boldsymbol{r} - \frac{\partial f}{\partial \boldsymbol{\chi}} \cdot \boldsymbol{h} \tag{36}$$

where the plastic flow direction $\boldsymbol{r}$ and the plastic moduli $\boldsymbol{h}$ are now characterized by

$$\left.\begin{aligned}
\boldsymbol{r} &= \frac{\partial g}{\partial \boldsymbol{\sigma}} \\
\boldsymbol{h} &= \frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{\epsilon}^{\mathrm{p}}} : \frac{\partial g}{\partial \boldsymbol{\sigma}}
\end{aligned}\right\} \tag{37}$$

To propel the iteration Eq. (35), a numerical evaluation of f is necessary, given that the explicit expression of f is generally absent. The value of f is determined by the z value, which corresponds to the scalar root of Eq. (29) when certain fixed $\boldsymbol{\sigma}$ and $\boldsymbol{\chi}$ are provided.

Further, the value of z can be bracketed in the interval $(0, z^{\max})$ if the stress state is located outside the yield surface, with $z^{\max}$ representing the maximum value of scaling factor that a user can specify. Under these circumstances, there is a range of root-finding algorithms available, such as the fast but non-guaranteed NR method, and the guaranteed Ridder's method and Brent's method (Press et al., 2007; Virtanen et al., 2020). Here, we propose a combined approach (see Fig. A1 in Appendix A) that offers both guarantees and sufficient computational efficiency for root approximation. In this study, the built-in optimize.brenth function in SciPy (Virtanen et al., 2020) is employed as the guaranteed module for a Python-based material point test (Harris et al., 2020). Additionally, the implementation of Brent's method given by Press (1996) is tailored for a user-defined subroutine (UMAT) in Abaqus/Standard (Smith, 2009) for finite element tests. The following parameter settings are employed when using the proposed root-finding method. The algorithm is tasked with seeking a numerical solution with an absolute error less than tolerance $TOL = 1 \times 10^{-8}$, provided that the iteration number is below the maximum limit of 50. The maximum value $z^{\max} = 50$ is assigned to z^{last} in the first computation of the scaling ratio in a plastic_corrector step.

The same procedure is preemptively employed for the

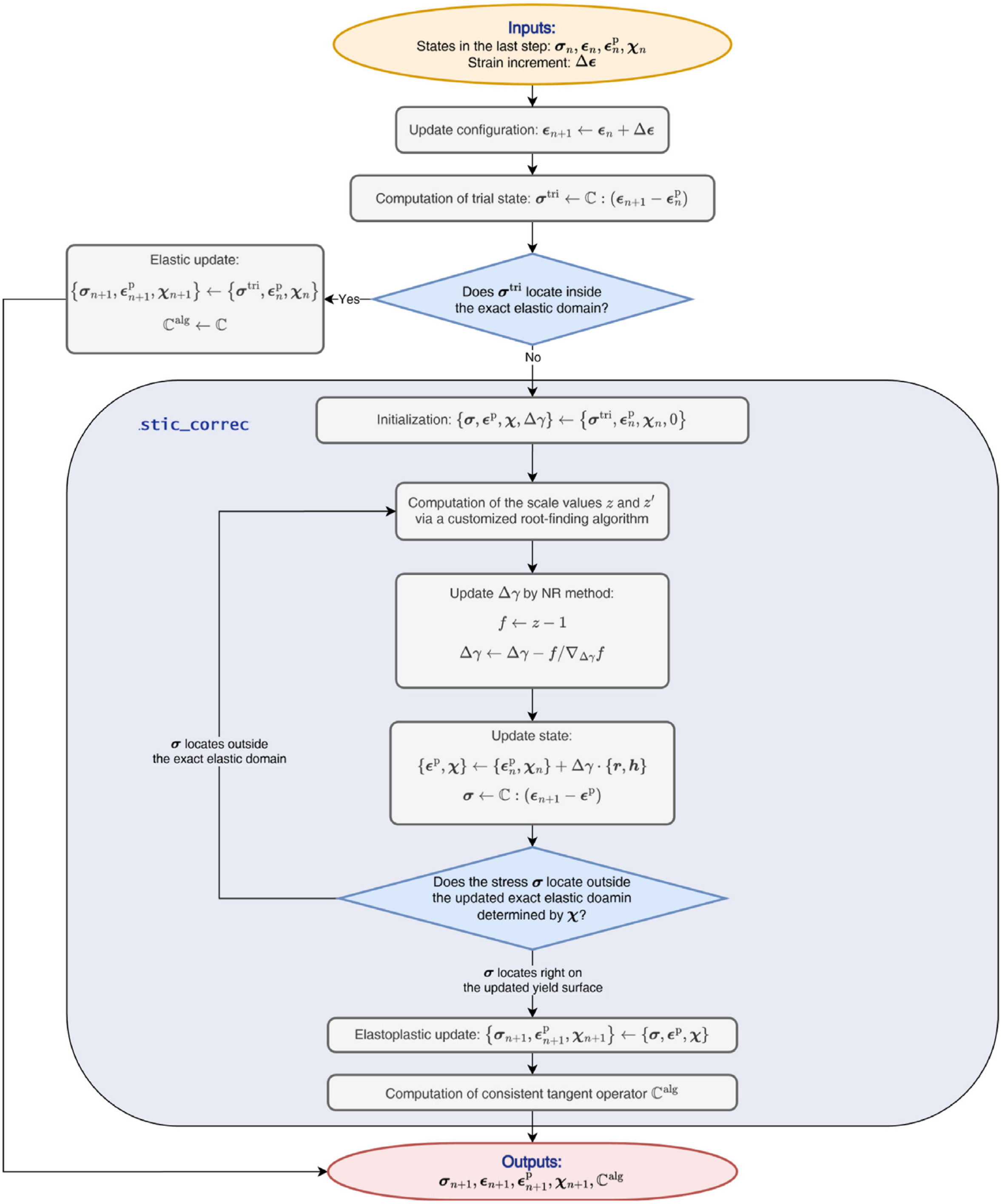


Fig. 3. Flowchart of the Hop-to-Hug (H2H) algorithm.

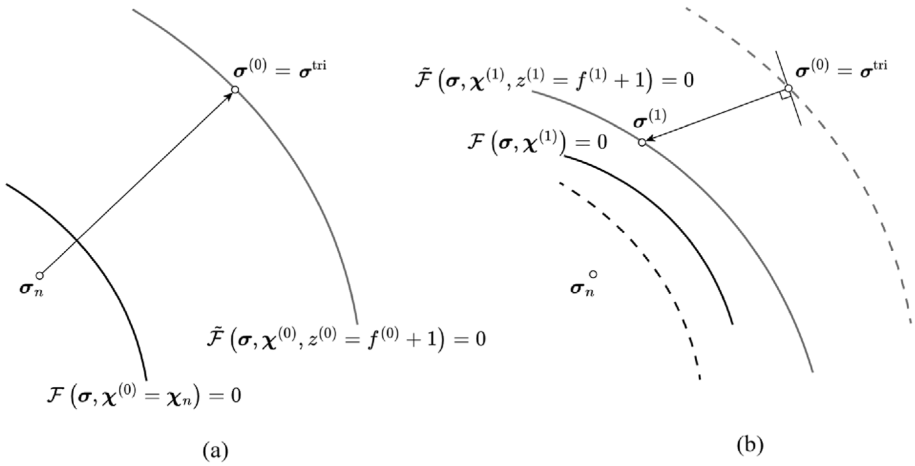


Fig. 4. Geometric interpretation of the plastic_corrector procedure in the Hop-to-Hug algorithm. (a) The elastic trial step, and (b) the first correction step.

evaluation of z' value for the plastic potential. As such, the method proposed here is summarized in Fig. 3. The core of this algorithm lies in the enhancement of the plastic_corrector, which improves the ability to deal with many potential hazards when the stress state is beyond the yield surface, through the construction of the quasiconvex yield function f. The notation $\mathbb{C}^{\mathrm{alg}}$ is the algorithmic stiffness tensor for global iteration, which will be elaborated in the subsequent section.

For the sake of understanding, Fig. 4 illustrates the plastic_corrector procedure in stress space. It can be observed that, in the H2H algorithm, stress states that have not yet converged (i.e. outside the intermediate yielding surface $\mathscr{F}[\boldsymbol{\sigma}, \boldsymbol{\chi}^{(i)}] = 0$ will form a homothetic yield surface that passes through this stress point. After a plastic correction step, the intermediate yield surface expands outward (hops outward), while the homothetic yield surface contracts inward (hops inward). Upon convergence of the plastic correction algorithm, the two yielding surfaces will coincide (hug) each other, and the stress state will be successfully corrected onto the intended yield surface. That is why the algorithm is named "Hop-to-Hug".

3.3. Consistent tangent operator

To address nonlinear IBVPs in solid mechanics using numerical methods such as the finite element method, it is conventionally necessary to apply external loads incrementally and obtain a new equilibrium step by deploying implicit iterative techniques, e.g. the Newton-Raphson method and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Liu and Nocedal, 1989). As usual, the new equilibrium state denoted by step $n+1$ is characterized by

$$\left\{\boldsymbol{F}(\boldsymbol{\sigma})_{n+1}\right\} = \left\{\boldsymbol{F}^{\mathrm{ext}}\right\}_{n+1} - \left\{\boldsymbol{F}^{\mathrm{int}}(\boldsymbol{\sigma})\right\}_{n+1} = \left\{\boldsymbol{0}\right\} \tag{38}$$

where $\{\boldsymbol{F}^{\mathrm{ext}}\}_{n+1}$ and $\{\boldsymbol{F}^{\mathrm{int}}(\boldsymbol{\sigma})\}_{n+1}$ are respectively the external force and internal force vector at the global level. For the strain-driving method, the following linearized formulation is adopted for the iterative process:

$$[\boldsymbol{K}]^{(i)}_{n+1}\{\delta\boldsymbol{U}\}^{(i+1)}_{n+1} = -\left[\left\{\boldsymbol{F}^{\mathrm{ext}}\right\}_{n+1} - \left\{\boldsymbol{F}^{int}(\sigma)\right\}^{(i)}_{n+1}\right] \tag{39}$$

where $\{\delta\boldsymbol{U}\}^{(i+1)}_{n+1}$ denotes the correction to incremental displacement vector at iteration $i+1$, and $[\boldsymbol{K}]^{(i)}_{n+1}$ is the global tangent stiffness matrix given by

$$[\boldsymbol{K}]^{(i)}_{n+1} = \left[\frac{\partial\boldsymbol{F}}{\partial\boldsymbol{\sigma}}\right]^{(i)}_{n+1}\left[\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{\epsilon}}\right]^{(i)}_{n+1}\left[\frac{\partial\boldsymbol{\epsilon}}{\partial\boldsymbol{U}}\right]^{(i)}_{n+1} \tag{40}$$

The tangent stiffness matrix $[\boldsymbol{K}]^{(i)}_{n+1}$ is assembled by combining the local stiffness matrices which reflect the local constitutive law. The local stiffness $[\partial\boldsymbol{\sigma}/\partial\boldsymbol{\epsilon}]^{\mathrm{loc},i}_{n+1}$ that maintains the consistent behavior of Eq. (39) is referred to as the consistent tangent operator (CTO). It is known that the CTO preserves the asymptotic rate of quadratic convergence in NR iteration, in contrast to the continuum tangent operator (CON) which only manages to deliver a slower convergence rate. In general, deriving the CTO requires the second-order derivatives of both the yield function and plastic potential, which is quite laborious for complex models. Furthermore, the explicit expression of CTO cannot be directly obtained from CPA results, even though a convergent state has been clearly determined. It is therefore that many constitutive models are willing to sacrifice a certain degree of convergence rate and opt for the CON for numerical computation. In order to complete the H2H algorithm, a numerical method for determining CTO in principal stress space is proposed. The method proposed here still inevitably requires the computation of second-order derivatives.

The spectral form of the stress tensor after the plastic_corrector process can be written as

$$\boldsymbol{\sigma} = \sum_{A=1}^{3}\sigma_A\boldsymbol{m}^{(A)},\ \boldsymbol{m}^{(A)} = \boldsymbol{n}^{(A)}\otimes\boldsymbol{n}^{(A)} \tag{41}$$

Therefore, the CTO by definition is given by

$$\mathbb{C}^{\mathrm{alg}} = \frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{\epsilon}} = \sum_{A=1}^{3}\sum_{B=1}^{3}a_{AB}\boldsymbol{m}^{(A)}\otimes\boldsymbol{m}^{(A)} + \sum_{A=1}^{3}\sigma_A\boldsymbol{\omega}^{(A)} \tag{42}$$

where $a_{AB} = \partial\sigma_A/\partial\epsilon_B$ originates from the constitutive response, and $\boldsymbol{\omega}^{(A)} = \partial\boldsymbol{m}^{(A)}/\partial\boldsymbol{\epsilon}$ indicates the rotation of principal directions. The determination of $\boldsymbol{\omega}^{(A)}$ is unrelated to the constitutive model adopted, for which the calculation method in the work (Itskov, 2000) is adopted.

With the motive of obtaining the term $(\partial\sigma_A/\partial\epsilon_B)$, a semi-analytical iterative approach is proposed based on the method covered by Starman et al. (2014). Considering the i th plastic_corrector iteration in the H2H algorithm, it holds

$$\left.\begin{aligned}&\Phi_1=\psi^{(i)}+\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\left[\{Y\}^{(i+1)}-\{Y\}^{(i)}\right]=0\\&\{\Phi_2\}=\{Y\}^{(i+1)}-\{Y\}^{(i)}+\delta(\Delta\gamma)^{(i)}\{\Theta\}^{(i)}=\{0\}\end{aligned}\right\}\tag{43}$$

where $\psi=\mathscr{F}$ or f determined by the yield function being activated. Assuming that $\boldsymbol{\chi}$ contains l parameters, the variable vector $\{Y\}$ and its derivative vector $\{\Theta\}$ (also determined by the plastic potential $\mathscr{G}$ or g being activated) are defined by

$$\left.\begin{aligned}&\{Y\}=\{\sigma_1,\sigma_2,\sigma_3,\chi_1,\chi_2,\cdots,\chi_l\}\\&\{\Theta\}=\frac{\partial\{Y\}}{\partial(\Delta\gamma)}\end{aligned}\right\}\tag{44}$$

The variations of equations Φ_1 and $\{\Phi_2\}$ should be zero:

$$\begin{aligned}&\mathrm{d}\Phi_1=\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\mathrm{d}\{Y\}^{(i)}+\left\{\frac{\partial^2\psi}{\partial\{Y\}\partial\{Y\}}^{(i)}\left[\{Y\}^{(i+1)}-\{Y\}^{(i)}\right]\right\}\cdot\mathrm{d}\{Y\}^{(i)}\\&+\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\left[\mathrm{d}\{Y\}^{(i+1)}-\mathrm{d}\{Y\}^{(i)}\right]=0\end{aligned}\tag{45}$$

$$\begin{aligned}&\mathrm{d}\{\Phi_2\}=\mathrm{d}\{Y\}^{(i+1)}-\mathrm{d}\{Y\}^{(i)}+\{\Theta\}^{(i)}\mathrm{d}\left[\delta(\Delta\gamma)^{(i)}\right]+\delta(\Delta\gamma)^{(i)}\\&\mathrm{d}\{\Theta\}^{(i)}=\{0\}\end{aligned}\tag{46}$$

Substituting Eq. (46) into Eq. (45) and taking advantage of

$$\frac{\partial\psi}{\partial\{Y\}}^{(i+1)}=\frac{\partial\psi}{\partial\{Y\}}^{(i)}+\frac{\partial^2\psi}{\partial\{Y\}\partial\{Y\}}^{(i)}\left[\{Y\}^{(i+1)}-\{Y\}^{(i)}\right]\tag{47}$$

One obtains that

$$\begin{aligned}&\mathrm{d}\left[\delta(\Delta\gamma)^{(i)}\right]=\left[\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\{\Theta\}^{(i)}\right]^{-1}\\&\cdot\left[\frac{\partial\psi}{\partial\{Y\}}^{(i+1)}\cdot\mathrm{d}\{Y\}^{(i)}-\delta(\Delta\gamma)^{(i)}\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\mathrm{d}\{\Theta\}^{(i)}\right]\end{aligned}\tag{48}$$

Taking the derivative of Eq. (48) with respect to the principal strain increment of the current global step, $\{\Delta\boldsymbol{\epsilon}\}=\{\Delta\epsilon_1,\Delta\epsilon_2,\Delta\epsilon_3\}$, it holds:

$$\begin{aligned}&\frac{\partial\delta(\Delta\gamma)^{(i)}}{\partial\{\Delta\boldsymbol{\epsilon}\}}=\left[\frac{\partial\psi}{\partial\{Y\}}^{(i)}\cdot\{\Theta\}^{(i)}\right]^{-1}\\&\cdot\left[\frac{\partial\psi}{\partial\{Y\}}^{(i+1)}-\delta(\Delta\gamma)^{(i)}\frac{\partial\psi}{\partial\{Y\}}^{(i)}\frac{\partial\{\Theta\}}{\partial\{Y\}}^{(i)}\right]\frac{\partial\{Y\}}{\partial\{\Delta\boldsymbol{\epsilon}\}}^{(i)}\end{aligned}\tag{49}$$

With Eq. (49) at hand, an explicit iterative equation can be extracted from Eq. (46):

$$\frac{\partial\{Y\}}{\partial\{\Delta\boldsymbol{\epsilon}\}}^{(i+1)}=\left\{[I_Y]-\delta(\Delta\gamma)^{(i)}\frac{\partial\{\Theta\}}{\partial\{Y\}}^{(i)}\right\}\frac{\partial\{Y\}}{\partial\{\Delta\boldsymbol{\epsilon}\}}^{(i)}-\{\Theta\}^{(i)}\otimes\frac{\partial\delta(\Delta\gamma)^{(i)}}{\partial\{\Delta\boldsymbol{\epsilon}\}}\tag{50}$$

where $[\boldsymbol{I}_Y]$ is an identity matrix. The iterative equation starts from initial value:

$$\frac{\partial\{Y\}}{\partial\{\Delta\boldsymbol{\epsilon}\}}^{(0)}=\begin{bmatrix}\lambda+2\mu&\lambda&\lambda\\\lambda&\lambda+2\mu&\lambda\\\lambda&\lambda&\lambda+2\mu\\0&0&0\\\vdots&\vdots&\vdots\\0&0&0\end{bmatrix}\tag{51}$$

where λ and μ are the Lamé constants.

Updating Eqs. (49) and (50) during the iterative process of the H2H method, the required term $\partial\sigma_A/\partial\epsilon_B$ can be acquired from the top 3×3 section of Eq. (50) when the plastic_corrector process converges. However, this iterative method inevitably involves the need to compute the second-order derivatives within $(\partial\{\Theta\}/\partial\{Y\})^{(i)}$. Therefore, the semi-analytical approach suggested by Starman et al. (2014) is adopted here, which involves using the forward finite difference method to approximate the term $(\partial\{\Theta\}/\partial\{Y\})^{(i)}$:

$$\left(\frac{\partial\{\Theta\}}{\partial\{Y\}}^{(i)}\right)^{y\text{th column}}=\frac{\{\Theta\}|_{\{Y\}^{(i)}+\upsilon_y\{\boldsymbol{e}_y\}}^{\ \ -\{\Theta\}|_{\{Y\}^{(i)}}}}{\upsilon_y}\tag{52}$$

where $\upsilon_y=|Y_y|\cdot\sqrt{\text{machine precision}}$ with Y_y is the y th component of $\{Y\}^{(i)}$; $\{\boldsymbol{e}_y\}$ is the unit vector of the same shape of $\{Y\}$, with its y th component equal to 1 and all other components equal to 0.

It is worth noting that although Eqs. (49) and (50) may appear complex, the components used in the equations, except for the additional computation required by Eq. (52), can be obtained from the H2H process and the equations themselves. This implies that acquiring the CTO using this method does not necessitate significant computational resources. Furthermore, this method eliminates the need for second-order derivatives of the yield function or plastic potential.

4. Numerical validation

The objective of this section is to illustrate the improved robustness of the proposed H2H method by means of numerical examples. Initially, a convergence analysis is conducted at a single material point for two models. Subsequently, a constitutive model with the H2H algorithm is implemented into the Abaqus/Standard platform via a user interface. Following this, a range of element-level tests and an analysis of a flexible strip footing analysis are performed.

4.1. Convergence analysis at the material point

A comprehensive convergence analysis is conducted at a single material point (integration point), for two classes of constitutive models: the Lade model and a modified version of the UDP model, because they collectively encompass the four types of numerical problems mentioned above, and thus possess sufficient representativeness. The investigation proceeds in accordance with the following procedure. Firstly, an initial yield surface is specified, after which the trial stress points are uniformly dispersed within a specific stress domain. Subsequently, the trial stress points are subjected to the plastic corrector using the CPA method and the H2H method, respectively. The number of iterations necessary for each point to attain a state of convergence is documented. All of these tasks are executed in the Python programming language, harnessing the powerful NumPy (Harris et al., 2020) and SciPy

(Virtanen et al., 2020) libraries.

4.1.1. Lade model

The Lade model, convexified by selecting the origin of HW stress space as the reference point, is subjected to two sets of convergence tests, with one set employing the linearly elastic-perfectly plastic model and the other utilizing a linearly elastic model with an exponential hardening law. The strain-hardening model takes η_1 as a single hardening parameter, which evolves as follows:

$$\eta_1 = \eta_m - (\eta_m - \eta_0)\exp(-\kappa_1 \bar{\epsilon}^{\mathrm{p}}) \tag{53}$$

where $\bar{\epsilon}^{\mathrm{p}}$ denotes the equivalent plastic strain.

Both cases adopt parameters $m^{\mathrm{L}} = 0.75$ and $a^{\mathrm{L}} = 10$, while the set without hardening takes constant $\eta_1 = 30$ and the set with hardening takes $\eta_0 = 30$, $\eta_m = 40$, and $\kappa_1 = 20$. Both analysis sets are conducted in the trial stress space of $(p^{\mathrm{tri}}, q^{\mathrm{tri}}) \in (-2,8) \times (0,6)$. Lode angles $\theta = -\pi/6$, 0, and $\pi/6$ are selected to show the influence of the third stress invariant. A total of 384,000 trial stress points are explored for each case. The results are presented in iteration-stress contour plots, as in Figs. 5 and 6 for models without and with hardening, respectively. For both cases, an absolute tolerance of $|\mathscr{F}| < TOL = 1 \times 10^{-8}$ for stress state returned to the desired branch of yield surface is adopted as the flag of convergence.

Both sets of analysis show that the H2H method outperforms the non-convexified model with the CPA method in terms of robustness, as the H2H method extends the range of convergence region of trial stress states. Due to its non-uniqueness and non-convexity, the Lade model exhibits significant non-convergence regions when the CPA method is employed. The H2H method overcomes the shortcomings of the Lade model and yields a greater region of convergence. However, one can observe that a certain region of non-convergence still exists when the H2H method is used, which can be attributed to the discontinuity of the first-order derivative at the apex in the Lade model.

4.1.2. Modified UDP model

In order to incorporate the effect of the Lode angle in the UDP model, this study introduces a shape function $\Gamma(\theta)$ into the original model to govern the shape of the yield surface on the π-plane. Therefore, the yield function of the modified UDP (MUDP) model is written as

$$\mathscr{F} = \left[\frac{q\Gamma(\theta)}{\sigma_0}\right]^2 - \left(1 + 3\alpha\frac{p}{\sigma_0}\right)^2 + 2\phi\left(1 + 3\alpha\frac{p}{\sigma_0}\right)\cosh\left[\frac{1}{2\alpha}\ln\left(1 + 3\alpha\frac{p}{\sigma_0}\right)\right] - \phi^2 \tag{54}$$

where $\Gamma(\theta)$ transforms the shape of yield surface in π-plane into a SMP style (Matsuoka, 1974) by

$$\Gamma(\theta) = \beta_1 \cos\left\{\frac{1}{3}\arccos(\beta_2 \sin 3\theta)\right\} \tag{55}$$

where $\beta_1 = 1/(3\alpha)$ and $\beta_2 = \sqrt{\alpha^2(3 - 4\alpha^2)^2}$.

The determination of parameters β_1 and β_2 is inspired by Lagioia and Panteghini (2016) and is also based on the micromechanics meaning of frictional constant α (Thoré et al., 2009; Durban et al., 2010). A representative yield surface is depicted in Fig. 7. It is worth noting that the modification of the original UDP model is more like a phenomenological treatment. A truly comprehensive consideration of the Lode angle effect on porous materials necessitates the attainment of consistent homogenization results from the Lode angle effect on the solid phase at the microscale, just as the works of Anoukou et al. (2016), Brach et al. (2018), and Fan et al. (2021). Nevertheless, these consistent products, characterized by their intricate mathematical expression, present challenges when attempting to conduct practical constitutive modeling. This justifies the simple modification given by Eqs. (54) and (55).

The MUDP model is also convexified by choosing the origin of HW stress space as the reference point. Three sets of convergence analysis tests are performed for the MUDP model: (1) a linearly elastic-perfectly plastic model, (2) a linear hardening model, and (3) a physical-based hardening model. The second model takes porosity ϕ as the single hardening parameter, which obeys the following law:

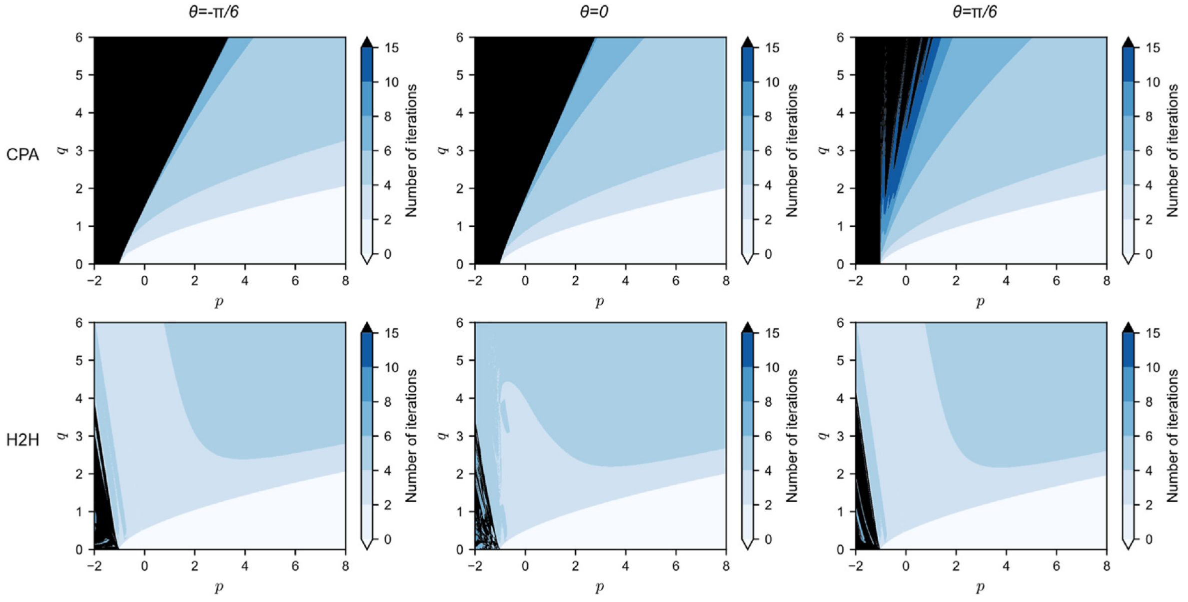


Fig. 5. Iteration-stress contour plots for linearly elastic-perfectly plastic Lade model. The upper and lower rows represent the computational results of the CPA method and the H2H method, respectively. The left, middle, and right columns correspond to the cases with Lode angles $\theta = -\pi/6$, 0, and $\pi/6$, respectively. The black area represents the region of non-convergence or regions that lead to program crashes.

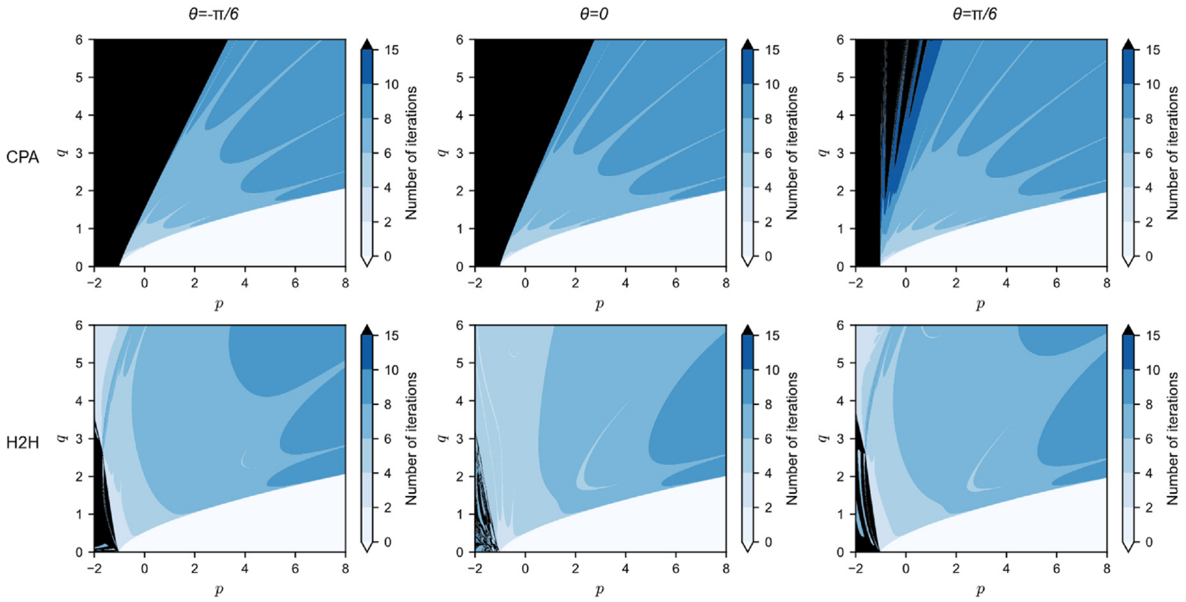


Fig. 6. Iteration-stress contour plots for the Lade model with linear elasticity and an exponential hardening law. The upper and lower rows represent the computational results of the CPA method and the H2H method, respectively. The left, middle, and right columns correspond to the cases with Lode angles $\theta = -\pi/6$, 0, and $\pi/6$, respectively. The black area represents the region of non-convergence or regions that lead to program crashes.

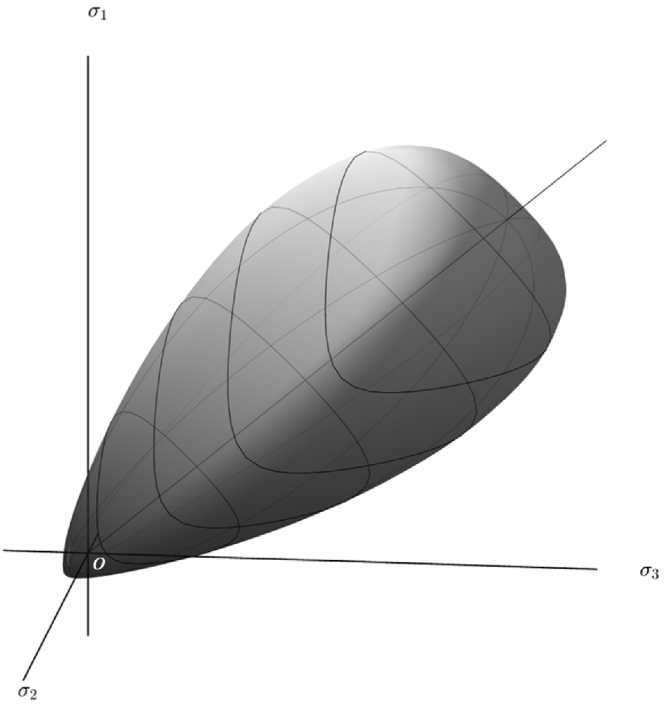


Fig. 7. A representative yield surface of the MUDP model in principal stress space. Parameters $\alpha = 0.3$, $\phi = 0.2$, and $\sigma_0 = 10$ MPa are adopted.

$$\phi = \phi_0 - \kappa_2 \bar{\epsilon}^{\mathrm{p}} \tag{56}$$

where ϕ_0 is the initial porosity and κ_2 is a material constant. The third model also takes ϕ as the hardening parameter, with a physical-based evolution law written as

$$\phi = 1 - (1 - \phi_0)\exp\left(3\alpha\bar{\epsilon}^{\mathrm{s,p}} - \epsilon_{\mathrm{v}}^{\mathrm{p}}\right) \tag{57}$$

where ϕ_0 is also the initial porosity, $\epsilon_{\mathrm{v}}^{\mathrm{p}}$ is the volumetric plastic strain of the porous material, and $\bar{\epsilon}^{\mathrm{s,p}}$ is the equivalent plastic strain of the solid phase of the porous material given by Gurson's energy condition (Gurson, 1977):

$$\bar{\epsilon}^{\mathrm{s,p}} = \frac{\boldsymbol{\sigma} : \boldsymbol{\epsilon}^{\mathrm{p}}}{(1 - \phi)\sigma_0} \tag{58}$$

All three cases adopt parameters $\alpha = 0.3$ and $\sigma_0 = 1$ MPa. The set without hardening takes constant porosity $\phi = 0.2$, the set with linear hardening takes $\phi_0 = 0.2$ and $\kappa_2 = 1$, and the set with physical-based hardening takes initial porosity $\phi_0 = 0.2$. A total of 400,000 trial stress states are explored in stress space $(p^{\mathrm{tri}}, q^{\mathrm{tri}}) \in (-5, 15) \times (0, 8)$ for each of the three selected Lode angles $\theta = -\pi/6$, 0, and $\pi/6$. All cases identify $|\mathscr{F}| < 1 \times 10^{-8}$ as the indicator to cease the plastic_corrector loop. Owing to its highly nonlinear hardening law, the third model requires more iteration to find the convergent solution than the first two models. The analysis results are plotted in iteration-stress contour as shown in Figs. 8–10.

As depicted by these contour plots, the incorporation of the H2H method enables the attainment of the return-mapping process even within the gradient-undefined region of the MUDP model. Moreover, a crucial aspect lies in the fact that the MUDP model exhibits a smooth yield surface (implying continuous first-order derivatives), which renders the H2H method to eliminate the potential for computational failure caused by trial stress increment towards any direction. Overall, the H2H method does not increase significantly the number of iterations in most regions in all three tests. In certain regions, it even reduces the iteration count. For some large incremental steps in tests of the linear hardening model and physics-based hardening model, however, a significant increase in iteration number can be observed in the results of the H2H method due to extra nonlinearity.

4.1.3. Discussion

This part presents the results of single-point stress integration tests for two Lade models and three MUDP models. Based on the graphical representations in Figs. 6–10 and the statistical summary presented in Table A1 in Appendix A, the following conclusions can

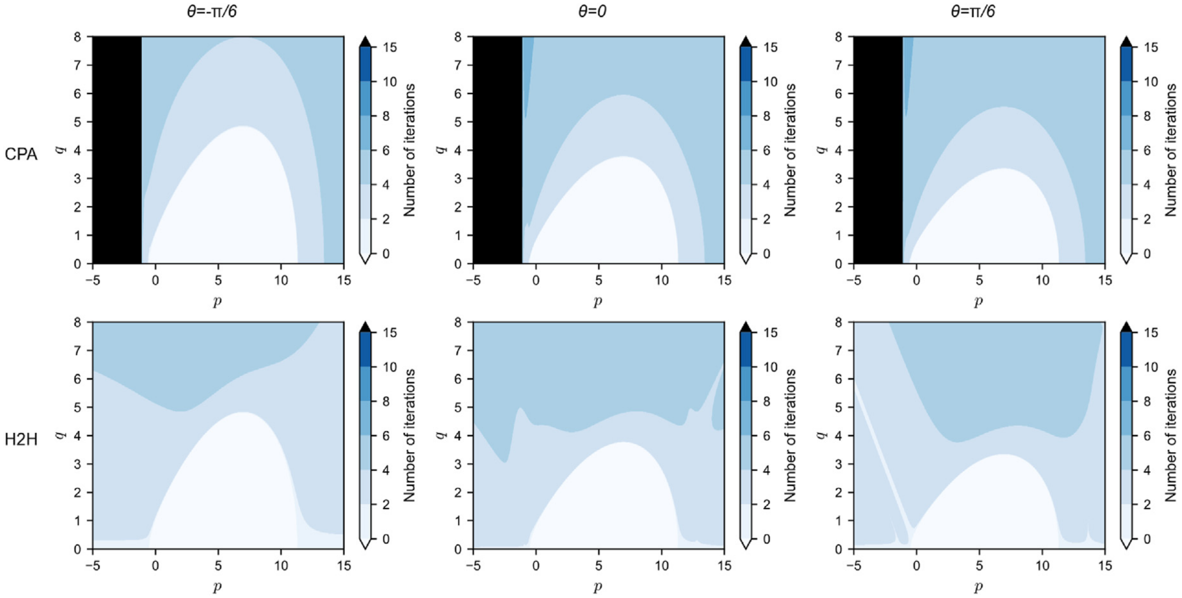


Fig. 8. Iteration-stress contour plots for linearly elastic-perfectly plastic MUDP model. The upper and lower rows represent the computational results of the CPA method and the H2H method, respectively. The left, middle, and right columns correspond to the cases with Lode angles $\theta = -\pi/6$, 0, and $\pi/6$, respectively. The black area represents the region of non-convergence or regions that lead to program crashes.

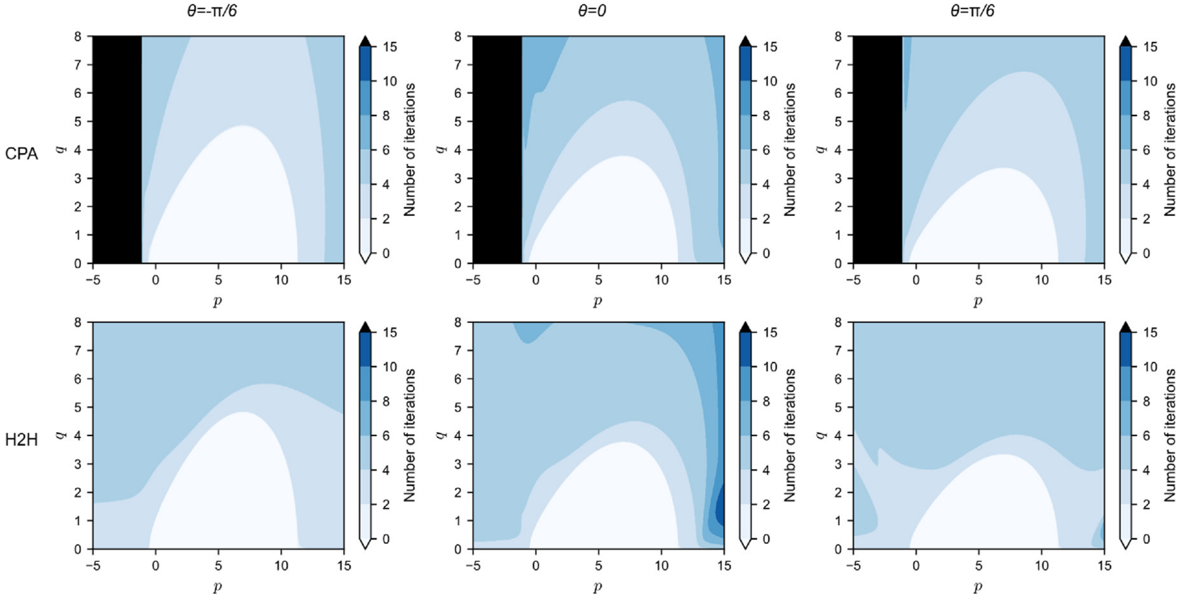


Fig. 9. Iteration-stress contour plots for the MUDP model with linear elasticity and a linear hardening law. The upper and lower rows represent the computational results of the CPA method and the H2H method, respectively. The left, middle, and right columns correspond to the cases with Lode angles $\theta = -\pi/6$, 0, and $\pi/6$, respectively. The black area represents the region of non-convergence or regions that lead to program crashes.

be drawn:

(1) The H2H method expands the CPA method's convergence region of trial stress space.
(2) The H2H method avoids the issue of incorrect return-mapping results caused by the non-convexity of the yield function and non-uniqueness of the elastic region.
(3) The H2H method makes the return-mapping (plastic_corrector) process feasible even within tentatively mathematically invalid trial stress regions.
(4) Though the H2H method generally increases the computational time for the plastic_corrector step (for the evaluation of z), the iteration number of the return-mapping process does not increase in most trial stress regions.
(5) The plastic_corrector steps of the H2H method require more computational time. The current Python-based analysis shows that the plastic_corrector step duration is typically 2–3 times longer than the correction step duration in conventional CPA methods. Nevertheless, the return-mapping process from most trial stress regions does not require more iterations compared to the CPA method, and the

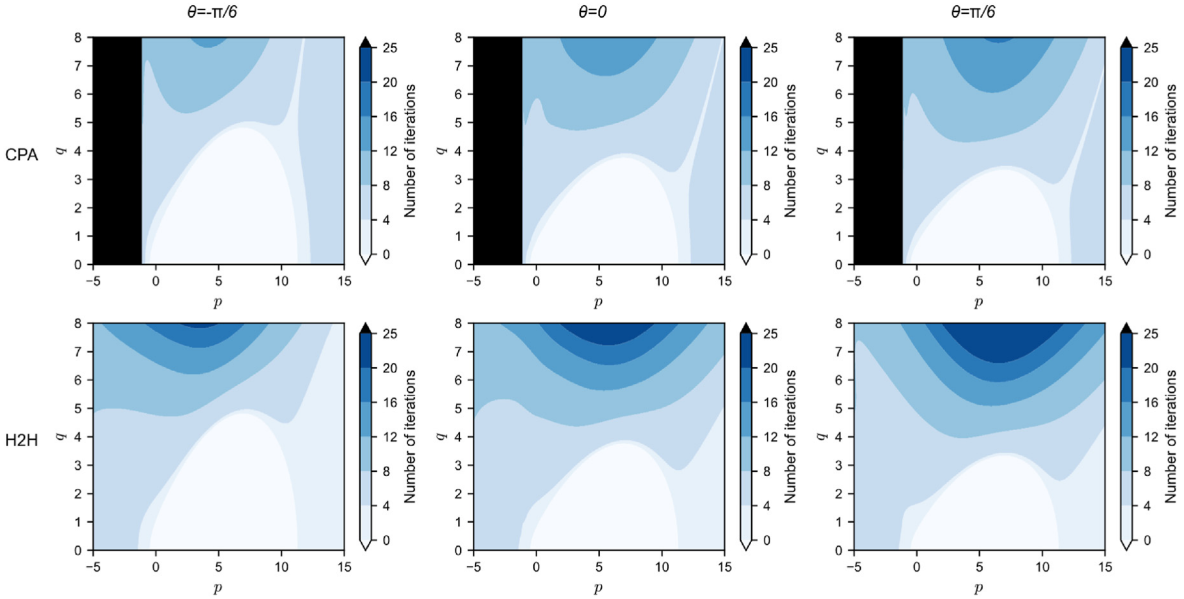


Fig. 10. Iteration-stress contour plots for the MUDP model with linear elasticity and a physical-based hardening law. The upper and lower rows represent the computational results of the CPA method and the H2H method, respectively. The left, middle, and right columns correspond to the cases with Lode angles $\theta = -\pi/6$, 0, and $\pi/6$, respectively. The black area represents the region of non-convergence or regions that lead to program crashes.

number of iterations is reduced in some trial stress regions. Therefore, there is a trade-off involved where, based on different requirements, one can choose a more robust but slower method, or opt for a faster method that may not necessarily be robust for even small incremental steps.

(6) It can be observed that the H2H method still exhibits local increases in the number of iterations due to some large incremental steps caused by the inherent instability of the Newton-Raphson method within the RM algorithm (Golchin et al., 2021). Certain measures that combine an accelerated Newton-Raphson method (McDougall and Wotherspoon, 2014; McDougall et al., 2019) with line search and other optimization techniques (Grippo et al., 1986; Nocedal and Wright, 1999) might be a potential approach to further address such issues.

In summary, the H2H algorithm, compared to the classical CPA method, is capable of enhancing the robustness of the return-mapping computations. Furthermore, one can observe that the procedure in the H2H algorithm to evaluate the scaling factor z and thereby form the quasi-convex yield function f is not necessary to activate for every incremental step. In other words, by conducting a more detailed analysis and classification of applied stress steps, the H2H method can be activated for those steps requiring robustness enhancement, while the classical CPA method can be employed for the steps that do not require it. The stress update algorithm formulated in this manner not only satisfies computational efficiency but also significantly enhances computational robustness.

4.2. Convergence analysis in finite element method

The MUDP model with a physics-based hardening law is implemented in a user-defined subroutine (UMAT) in Abaqus/Standard for finite element tests. The following analyses were conducted using the 8-node brick element (C3D8) in Abaqus/Standard (Smith, 2009). The hardware platform used consists of an Intel Core i7-9750H processor @ 2.60 GHz and 16 GB of RAM. The computer operates on a 64-bit Windows 11 operating system.

4.2.1. Tests on one element

In this example, the constitutive model MUDP with physical-based hardening is adopted with elastic constants $k = 12.5$ GPa and $\mu = 11.5$ GPa and parameters $\alpha = 0.33$, $\phi_0 = 0.25$, and $\sigma_0 = 15$ MPa to describe some rock-like materials. One C3D8 cubic element measuring 1 m × 1 m × 1 m, as shown in Fig. 11a, has the normal displacements constrained on its three adjacent orthogonal faces. Five monotonic displacement control loading paths are applied:

(1) A uniaxial tension test with the ultimate axial strain is set to $\epsilon_{11}^{\max} = 0.005$.

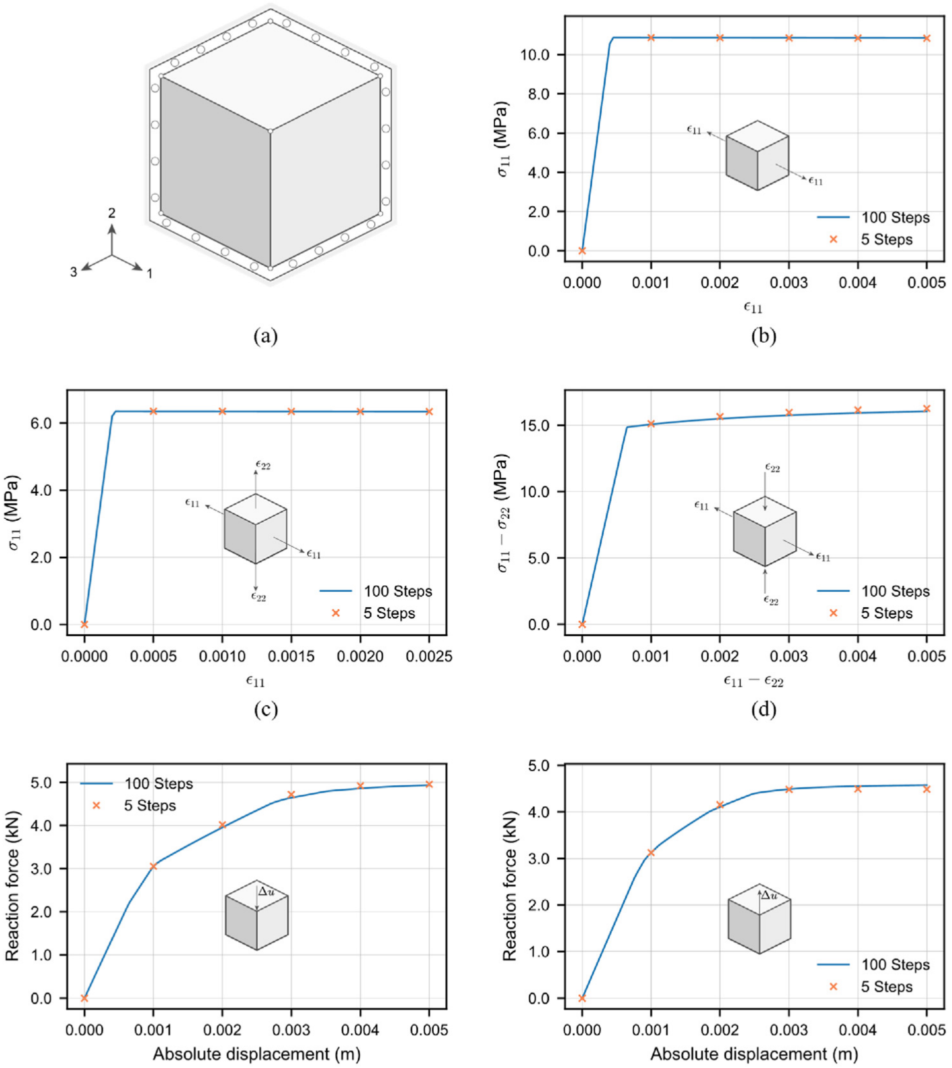
(2) A biaxial tension test with equally proportional stretching in 1-direction and 2-direction is used until reaching the ultimate strain state of $\epsilon_{11} = \epsilon_{22} = 0.0025$.

(3) A biaxial tension-compression test with equally proportional stretching in 1-direction and compression in 2-direction is used, resulting in the ultimate strain state of $\epsilon_{11} = -\epsilon_{22} = 0.0025$.

(4) A single nodal loading is applied to the free node with controlled compression in 2-direction, leading to an ultimate absolute nodal displacement of $\Delta u = 0.005$ m in inverse 2-direction.

(5) A single nodal loading is applied to the free node with controlled tension in 2-direction, resulting in an ultimate absolute nodal displacement of $\Delta u = 0.005$ m in 2-direction. All five loading paths are subjected to two different computation schemes: 100 incremental steps and 5 incremental steps, respectively, for each loading path.

The calculated results are presented in Fig. 11b, c, and d as the corresponding stress-strain curves, and in Fig. 11e and f as nodal reaction force-nodal absolute displacement curves of the free node in 2-direction. The solid blue line represents the results obtained using 100 incremental steps, while the orange crosses depict the results obtained using 5 incremental steps. The results of the large increment step (5 steps) simulation are almost consistent with those of the small increment step (100 steps). Compared to the results of small incremental steps, the final absolute relative errors



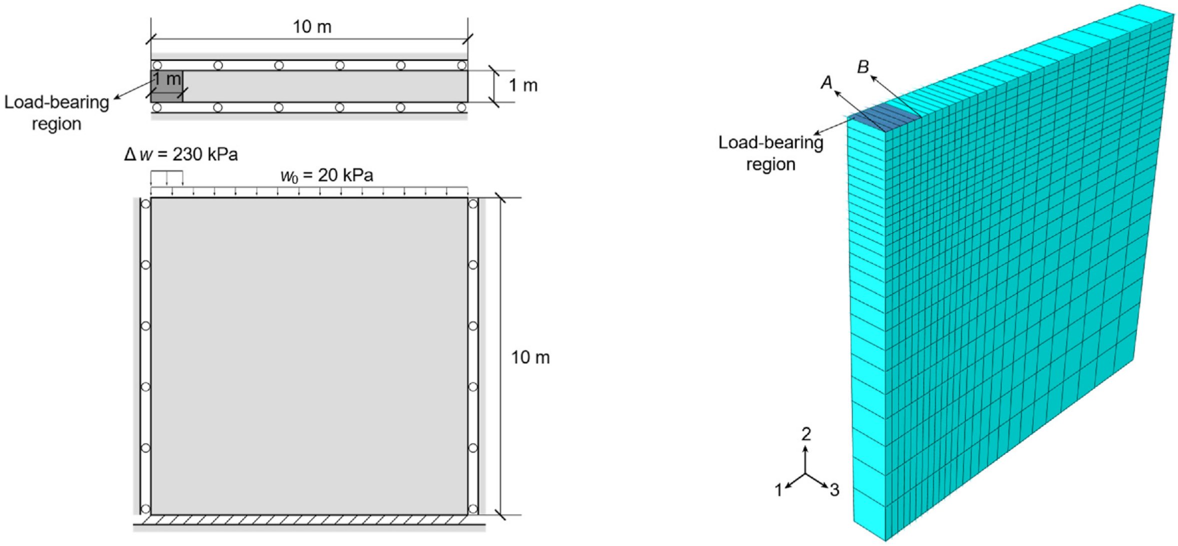
(a) (b) (c) (d) (e) (f)

Fig. 11. Element loading tests. (a) Illustration of the tested element with the prescribed boundary conditions before loading, (b) uniaxial tension test, (c) biaxial tension test, (d) biaxial tension-compression test, (e) single nodal compression loading test, and (f) single nodal tension loading test.

at the end of the loading stage in tests (1)–(5) (corresponding to Fig. 11b–f) are 1.22%, 0.05%, 1.31%, 0.40%, and 1.92%, respectively, for the large incremental step loading. This indicates that the H2H algorithm has sufficient convergence dealing with large step loading in the finite element method. Moreover, the stable performance of the H2H algorithm in the uniaxial and biaxial tensile tests is particularly encouraging. This is because large loading steps in the tensile direction mean that the trial stresses will directly fall into the mathematically invalid region of the MUDP model. If the traditional RM algorithm were used, the computation would be interrupted. During large-scale computational tasks, the simulation may not be able to proceed merely because the trial stresses of a few elements fall into the invalid region, which would be unacceptable. As analyzed in the previous section, there may be other types of numerical traps inherent in many complex and advanced, yet not fully evaluated, constitutive models. While the common choice to ensure the effective operation of the program is to select a simple and robust constitutive model for large-scale simulations (e.g. Sepehri et al., 2017), researchers and engineers may seek more sophisticated constitutive models to obtain more accurate simulation results, as long as the computation remains stable and other conditions permit. Compared to the direct use of the RM algorithm, the robustness exhibited by the H2H algorithm provides new possibilities for the implementation of more advanced constitutive models.

4.2.2. *Flexible strip footing analysis*

The strip footing problem is a common benchmark problem for the stress integration algorithm. In the strip footing problem, the emergence of singularity at the footing edge and the rotation of principal stresses provides a rigorous testing environment to check the robustness and convergence of numerical integration schemes in handling complex loading paths. To access the convergence ability of the H2H algorithm, a simple flexible strip footing problem is modeled with the MUDP model taking elastic constants $k = 5$ MPa and $\mu = 1.875$ MPa and parameters $\alpha = 0.25$, $\phi_0 = 0.2$, and $\sigma_0 = 0.06$ MPa to describe the mechanical behavior of the simulated soil foundation. The weight of the soil is set to 20 kN/m^3. The symmetric footing model is created and meshed as illustrated in Fig. 12. A total of 676 C3D8 elements with 1458 nodes are employed to simulate the plane strain condition by constraining the displacements in 3-direction. The model was subjected to loads in two stages. Initially, the self-weight and uniformly distributed surface load w_0 are applied with a single step. Then a uniformly distributed



(a) (b)

Fig. 12. Geometric illustration of the symmetrical flexible strip footing: The deep grey area in (a) and the steel blue area in (b) represent the load-bearing regions of the foundation. Points *A* and *B* are selected as analysis points for settlement assessment.

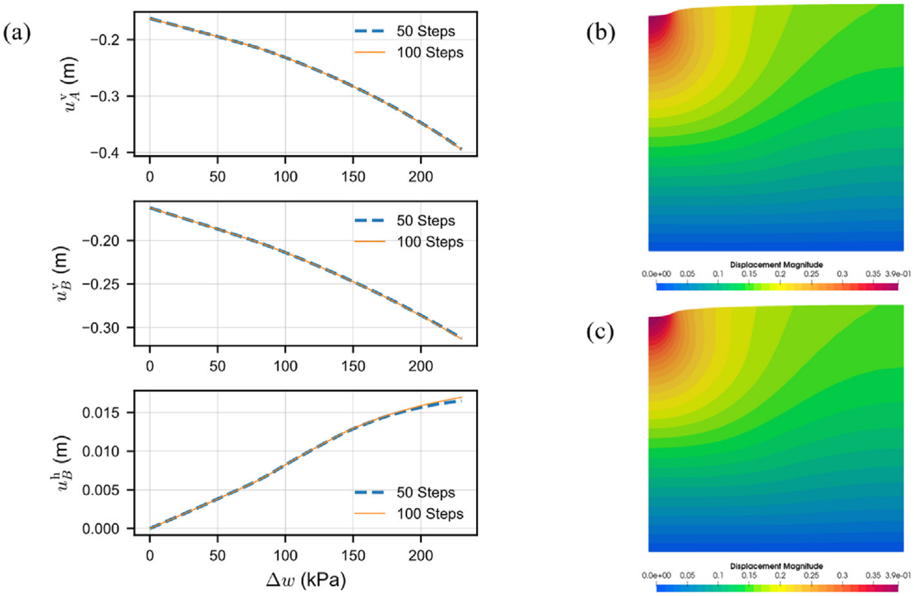


Fig. 13. (a) Comparisons between the modeling results of 50- and 100-step loading schemes utilizing the H2H algorithm, (b) displacement magnitude contour plot of 50-step loading scheme, and (c) displacement magnitude contour plot of 100-step loading scheme. The deformations in (b) and (c) are both magnified by factor 2. u_A^{v} represents the vertical displacement of point *A*, and u_B^{v} and u_B^{h} denote respectively the vertical and horizontal displacement of point *B*.

Table 2
Numerical solution of the displacements of points *A* and *B* with 50 and 100-step loading schemes.

Location	Displacement (m)	
	50 steps	100 steps
Vertical displacement of point *A*	−0.39400	−0.39458
Vertical displacement of point *B*	−0.31223	−0.31313
Horizontal displacement of point *B*	0.01649	0.01695

load Δ*w* generated by the flexible footing is incrementally imposed on the 1 m × 1 m area with 50 equal increments and 100 equal increments schemes, respectively (see Fig. 12).

The displacement data of points *A* and *B* (as indicated in Fig. 12b) and total settlement results are recorded and summarized in Fig. 13. The final settlement solutions in points *A* and *B* with two loading schemes are presented in Table 2. Although the finite element mesh used is relatively coarse, the displacement solutions for both 50 incremental steps and 100 incremental steps are nearly identical. The computation durations in the present hardware platform for the two schemes were recorded as 22 s and 43 s, respectively. As a comparison, using the built-in elastic-perfectly plastic Mohr-Coulomb model in Abaqus/Standard for the 100 incremental steps simulation costs around 24 s. The analysis results indicate that the H2H algorithm can handle complex loading conditions with acceptable computational cost and provide sufficient accuracy in the finite element method.

5. Conclusions

This study aims to address the numerical challenges encountered when solving nonlinear initial boundary value problems (IBVPs) with a particular focus on the stress integration or stress update method. The following conclusions are drawn:

(1) We begin with an overview of the widely recognized return-mapping algorithm family, also known as the elastic predictor-plastic corrector method, with a specific emphasis on the cutting-plane algorithm (CPA). Subsequently, we delve into the numerical issues that arise due to the mathematical properties of the yield function when $\mathscr{F} > 0$. These numerical challenges, including non-uniqueness, non-convexity, undefined gradients, and high curvature, are illustrated via examinations over four typical yield functions.
(2) To overcome these obstacles, we propose a robust stress integration method called the Hop-to-Hug (H2H) algorithm. The H2H algorithm is initially established based on the convexification of the general yield function, followed by the implementation of the elastic predictor-plastic corrector method. Additionally, a semi-analytical consistent tangent operator (CTO) for the H2H algorithm is proposed.
(3) A series of numerical experiments are conducted to validate the effectiveness of the H2H algorithm. These experiments encompass various levels of analysis, including a single integration point, a single finite element, and a structural IBVP. The results of the numerical validations demonstrate the robustness and convergence capabilities of the H2H algorithm.

Although the proposed method in this paper sustains a slightly higher computational cost compared to the CPA method, it offers enhanced stability in the implicit numerical process. Furthermore, this approach can be readily extended to address most stress integration computations related to plasticity. Future work will focus on the extension to non-smooth models as well as elasto-plastic damage models. Additionally, the accelerated Newton-Raphson method combined with some optimization techniques will be considered to enhance the stability and convergence rate of current plastic correction computation. All in all, the H2H algorithm presents an implementable and promising approach for enhancing implicit stress integration algorithms.

CRediT authorship contribution statement

Yanbin Chen: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Yuanming Lai:** Resources, Supervision, Writing – review & editing. **Enlong Liu:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 12372376 and U22A20596).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrmge.2024.07.009.

References

Anandarajah, A., 2010. Computational Methods in Elasticity and Plasticity. Springer New York, New York, NY.
Anoukou, K., Pastor, F., Dufrenoy, P., Kondo, D., 2016. Limit analysis and homogenization of porous materials with Mohr–Coulomb matrix. Part I: theoretical formulation. J Mech Phys Solids 91, 145–171.
Areias, P., Dias-da-Costa, D., Pires, E.B., Barbosa, J.I., 2012. A new semi-implicit formulation for multiple-surface flow rules in multiplicative plasticity. Comput. Mech. 49, 545–564.
Barlat, F., Aretz, H., Yoon, J.W., Karabin, M.E., Brem, J.C., Dick, R.E., 2005. Linear transfomation-based anisotropic yield functions. Int J Plasticity 21, 1009–1039.
Bigoni, D., Piccolroaz, A., 2004. Yield criteria for quasibrittle and frictional materials. Int J Solids Struct 41, 2855–2878.
Borja, R.I., Lee, S.R., 1990. Cam-Clay plasticity, Part 1: implicit integration of elasto-plastic constitutive relations. Comput Method Appl M 78, 49–72.
Borja, R.I., Lin, C.-H., Montáns, F.J., 2001. Cam-clay plasticity, Part IV: implicit integration of anisotropic bounding surface model with nonlinear hyperelasticity and ellipsoidal loading function. Comput Method Appl M 190, 3293–3323.
de Borst, R., Crisfield, M.A., Remmers, J.J., Verhoosel, C.V., 2012. Nonlinear Finite Element Analysis of Solids and Structures. John Wiley & Sons.
Boyd, S.P., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.
Brach, S., Anoukou, K., Kondo, D., Vairo, G., 2018. Limit analysis and homogenization of nanoporous materials with a general isotropic plastic matrix. Int J Plasticity 105, 24–61.
Bryant, E.C., Sun, W., 2019. A micromorphically regularized Cam-clay model for capturing size-dependent anisotropy of geomaterials. Comput Method Appl M 354, 56–95.
Cardoso, R.P.R., Yoon, J.W., 2009. Stress integration method for a nonlinear kinematic/isotropic hardening model and its characterization based on polycrystal plasticity. Int J Plasticity 25, 1684–1710.
Coombs, W.M., Crouch, R.S., 2011. Algorithmic issues for three-invariant hyperplastic Critical State models. Comput Method Appl M 200, 2297–2318.
Cornejo, A., Jiménez, S., Barbu, L.G., Oller, S., Oñate, E., 2022. A unified non-linear energy dissipation-based plastic-damage model for cyclic loading. Comput Method Appl M 400, 115543.
Coxeter, H.S.M., 1969. Introduction to Geometry. John Wiley & Sons, Inc.
Dafalias, Y.F., Manzari, M.T., 2004. Simple plasticity sand model accounting for fabric change effects. J. Eng. Mech. 130, 622–634.
Derbin, Y., Walker, J., Wanatowski, D., Marshall, A.M., 2018. Implementation of advanced constitutive models for the prediction of surface subsidence after underground mineral extraction. In: Wu, W., Yu, H.-S. (Eds.), Proceedings of China-Europe Conference on Geotechnical Engineering. Springer International Publishing, Cham, pp. 320–323.
Ding, Y., Huang, W., Sheng, D., Sloan, S.W., 2015. Numerical study on finite element implementation of hypoplastic models. Comput. Geotech. 68, 78–90.
Durban, D., Cohen, T., Hollander, Y., 2010. Plastic response of porous solids with pressure sensitive matrix. Mech. Res. Commun. 37, 636–641.
Fan, J., Brassart, L., Shen, W., Ge, X., 2021. An explicit formulation of the macroscopic strength criterion for porous media with pressure and Lode angle dependent matrix under axisymmetric loading. J. Rock Mech. Geotech. Eng. 13, 820–832.
Feng, W., Niu, S., Qiao, C., Zou, D., 2024. Nonlinear constitutive models of rock structural plane and their applications. J. Rock Mech. Geotech. Eng. 16, 790–806.
Foster, C.D., Regueiro, R.A., Fossum, A.F., Borja, R.I., 2005. Implicit numerical integration of a three-invariant, isotropic/kinematic hardening cap plasticity model for geomaterials. Comput Method Appl M 194, 5109–5138.
Golchin, A., Vardon, P.J., Coombs, W.M., Hicks, M.A., 2021. A flexible and robust yield function for geomaterials. Comput Method Appl M 387, 114162.
Grippo, L., Lampariello, F., Lucidi, S., 1986. A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. 23, 707–716.
Gurson, A.L., 1977. Continuum theory of ductile rupture by void nucleation and growth: Part I—yield criteria and flow rules for porous ductile media. J. Eng. Mater. Technol. 99, 297–300.
Hall, B.C., 2015. Lie Groups, Lie Algebras, and Representations: an Elementary Introduction, vol. 222. Springer International Publishing, Cham.
Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with NumPy. Nature 585, 357–362.
Hill, R., 1998. The Mathematical Theory of Plasticity. Oxford University Press, Oxford, New York.
Hill, R., 1948. A theory of the yielding and plastic flow of anisotropic metals. Proceedings of the Royal Society of London Series A, Math Phys Sci 193, 281–297.
Itskov, M., 2000. On the theory of fourth-order tensors and their applications in

computational mechanics. Comput Method Appl M 189, 419–438.

Jiao, H., Lv, Y., Chen, D., Huang, W., Su, Y., 2024. Numerical implementation of the hypoplastic model for SPH analysis of soil structure development in extremely large deformation. Comput. Geotech. 166, 106014.

Karaoulanis, F.E., 2013. Implicit numerical integration of nonsmooth multisurface yield criteria in the principal stress space. Arch Comput Method E 20, 263–308.

Khosravifar, A., Elgamal, A., Lu, J., Li, J., 2018. A 3D model for earthquake-induced liquefaction triggering and post-liquefaction response. Soil Dyn Earthq Eng 110, 43–52.

Lade, P.V., 1977. Elasto-plastic stress-strain theory for cohesionless soil with curved yield surfaces. Int J Solids Struct 13, 1019–1035.

Lagioia, R., Panteghini, A., 2016. On the existence of a unique class of yield and failure criteria comprising Tresca, von Mises, Drucker–Prager, Mohr–Coulomb, Galileo–Rankine, Matsuoka–Nakai and Lade–Duncan. P Roy Soc A-Math Phy 472, 20150713.

Lee, J., Bong, H.J., Lee, M.-G., 2021. Return mapping with a line search method for integrating stress of the distortional hardening law with differential softening. Comput. Struct. 257, 106652.

Lin, C., Li, Y.-M., 2015. A return mapping algorithm for unified strength theory model. Int J Numer Meth Eng 104, 749–766.

Liu, D.C., Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. Math. Program. 45, 503–528.

Liu, H., Abell, J., Diambra, A., Pisanò, F., 2018. Modelling the cyclic ratcheting of sands through memory-enhanced bounding surface plasticity. Geotechnique 69, 1–57.

Liu, H., Kementzetzidis, E., Abell, J.A., Pisanò, F., 2022. From cyclic sand ratcheting to tilt accumulation of offshore monopiles: 3D FE modelling using SANISAND-MS. Geotechnique 72, 753–768.

Lloret-Cabot, M., Wheeler, S.J., Gens, A., Sloan, S.W., 2021. Numerical integration of an elasto-plastic critical state model for soils under unsaturated conditions. Comput. Geotech. 137, 104299.

Matsuoka, H., 1974. Stress-strain relationships of sands based on the mobilized plane. Soils Found. 14, 47–61.

Matsuoka, H., Nakai, T., 1985. Relationship among tresca, mises, mohr-coulomb and matsuoka-nakai failure criteria. Soils Found. 25, 123–128.

McDougall, T.J., Wotherspoon, S.J., 2014. A simple modification of Newton's method to achieve convergence of order $1+\sqrt{2}$. Appl. Math. Lett. 29, 20–25.

McDougall, T.J., Wotherspoon, S.J., Barker, P.M., 2019. An accelerated version of Newton's method with convergence order $\sqrt{3}+1$. Results Appl Math 4, 100078.

Meserve, B.E., 2014. Fundamental Concepts of Geometry. Courier Corporation.

Nocedal, J., Wright, S.J., 1999. Numerical Optimization. Springer-Verlag, New York.

Ortiz, M., Simo, J.C., 1986. An analysis of a new class of integration algorithms for elastoplastic constitutive relations. Int J Numer Meth Eng 23, 353–366.

Panteghini, A., Lagioia, R., 2018. An approach for providing quasi-convexity to yield functions and a generalized implicit integration scheme for isotropic constitutive models based on 2 unknowns. Int J Numer Anal Met 42, 829–855.

Press, W.H. (Ed.), 1996. FORTRAN Numerical Recipes, second ed. Cambridge University Press, Cambridge.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. In: The Art of Scientific Computing, third ed. Cambridge University Press, Cambridge.

Ren, L., Zhao, L.-Y., Niu, F.-J., 2022. A physically-based elastoplastic damage model for quasi-brittle geomaterials under freeze-thaw cycles and loading. Appl. Math. Model. 106, 276–298.

Roscoe, K.H., Burland, J., 1968. On the Generalized Stress-Strain Behavior of Wet Clays. Cambridge University Press, Cambridge.

Scalet, G., Auricchio, F., 2018. Computational methods for elastoplasticity: an overview of conventional and less-conventional approaches. Arch Computat Methods Eng 25, 545–589.

Schofield, A.N., Wroth, P., 1968. Critical State Soil Mechanics. McGraw-Hill, New York.

Sepehri, M., Apel, D.B., Hall, R.A., 2017. Prediction of mining-induced surface subsidence and ground movements at a Canadian diamond mine using an elastoplastic finite element model. Int. J. Rock Mech. Min. Sci. 100, 73–82.

Simo, J.C., Hughes, T.J.R., 1998. Computational Inelasticity. Springer, New York.

Simo, J.C., Ortiz, M., 1985. A unified approach to finite deformation elastoplastic analysis based on the use of hyperelastic constitutive equations. Comput Method Appl M 49, 221–245.

Simo, J.C., Taylor, R.L., 1985. Consistent tangent operators for rate-independent elastoplasticity. Comput Method Appl M 48, 101–118.

Sloan, S.W., 1987. Substepping schemes for the numerical integration of elasto-plastic stress–strain relations. Int J Numer Meth Eng 24, 893–911.

Sloan, S.W., Abbo, A.J., Sheng, D., 2001. Refined explicit integration of elastoplastic models with automatic error control. Eng. Comput. 18, 121–194.

Smith, M., 2009. ABAQUS/standard User's Manual, Version 6.9. Dassault Systèmes Simulia Corp.

Starman, B., Halilovič, M., Vrh, M., Štok, B., 2014. Consistent tangent operator for cutting-plane algorithm of elasto-plasticity. Comput Method Appl M 272, 214–232.

Suryasentana, S.K., Burd, H.J., Byrne, B.W., Shonberg, A., 2021. Automated procedure to derive convex failure envelope formulations for circular surface foundations under six degrees of freedom loading. Comput. Geotech. 137, 104174.

Tamagnini, C., Castellanza, R., Nova, R., 2002. 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 Met 26, 963–1004.

Thoré, P., Pastor, F., Pastor, J., Kondo, D., 2009. Closed-form solutions for the hollow sphere model with Coulomb and Drucker–Prager materials under isotropic loadings. Cr Mecanique 337, 260–267.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., HarrisCR, Archibald AM., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272.

Wei, S.-M., Yuan, L., Cui, Z.-D., 2021. Application of closest point projection method to unified hardening model. Comput. Geotech. 133, 104064.

Xiao, Y., He, M., Qiao, Y., Cai, W., Jiang, Y., Zhu, H., 2024. A novel implementation method of GZZ-based constitutive model into FLAC3D. Tunn Undergr Sp Tech 145, 105601.

Zhao, L.-Y., Liu, L.-H., Lai, Y.-M., 2023. An enhanced cutting plane algorithm of elastoplastic constitutive models for geomaterials. Comput. Geotech. 155, 105253.

Zhao, L.-Y., Zhu, Q.-Z., Shao, J.-F., 2018. A micro-mechanics based plastic damage model for quasi-brittle materials under a large range of compressive stress. Int J Plasticity 100, 156–176.

Zhou, X., Lu, D., Su, C., Gao, Z., Du, X., 2022. An unconstrained stress updating algorithm with the line search method for elastoplastic soil models. Comput. Geotech. 143, 104592.

Zhu, Q.Z., Zhao, L.Y., Shao, J.F., 2016. Analytical and numerical analysis of frictional damage in quasi brittle materials. J Mech Phys Solids 92, 137–163.

Yanbin Chen obtained his BSc in civil engineering from Sichuan University, China in 2020. He is currently a PhD candidate in Safety Science and Disaster Reduction at Institute for Disaster Management and Reconstruction (IDMR), Sichuan University-The Hong Kong Polytechnic University, Chengdu, China. His research focuses on the constitutive modeling of geomaterials and other solid materials, including analytical modeling from a micromechanics perspective, development and improvement of stress integration algorithms, as well as application of advanced constitutive models for simulation purposes.