RESEARCH PAPER

Numerical integration and FE implementation of a hypoplastic constitutive model

Shun Wang[1] · Wei Wu[1] · Chong Peng[1] · Xuzhen He[1] · Deshan Cui[2]

Abstract

Hypoplastic constitutive equation based on nonlinear tensor functions possesses a failure surface but no yield surface. In this paper, we consider the numerical integration and FE implementation of a simple hypoplastic constitutive equation. The accuracy of several integration methods, including implicit and explicit methods, is examined by performing a set of triaxial compression tests. Adaptive explicit schemes show the best performance. In addition, the stress drift away from the failure surface is corrected with a predictor-corrector scheme, which is verified by two boundary value problems, i.e. rigid footing tests and slope stability.

1 Introduction

Hypoplasticity represents a class of incrementally nonlinear constitutive models [9, 11, 16, 17, 28, 42]. Unlike elastic–plastic models, there is no need to decompose the deformation into elastic and plastic deformations. Moreover, hypoplastic models do not make use of yield and potential surfaces [37]. Hypoplastic models are characterised by simple formulation and few material parameters and have been used to simulate soil behaviour in element tests and to solve boundary value problems with FEM [14, 25].

The performance of FEM depends on the efficiency of the numerical integration of constitutive equation. Recently, numerical integration of hypoplastic constitutive equations in finite element analysis is a topic of considerable interest. The fact that hypoplastic model has a single equation makes the implementation more straightforward than elastoplastic models [4, 30]. Since hypoplastic model does not have a yield surface, the stress return mapping

algorithms common for elastoplastic models are not needed. Research on the finite element implementation of hypoplastic models can be traced back to the early work by Sikora and Gudehus [26], in which a simple explicit forward Euler scheme with constant step sizes was adopted. Later, this method was used to investigate shear band formation in granular materials by Tejchman and Wu [31]. Roddeman [24] used a generalised midpoint algorithm and Heeres and de Borst [8] considered an implicit integration method together with Newton–Raphson iterative scheme for the stress integration of a hypoplastic model [32]. Tamagnini et al. [30] studied the accuracy of some explicit methods and the generalised midpoint algorithm. Recently, Ding et al. [4] showed that explicit methods with substepping and error control are suitable for the numerical implementation of hypoplastic models.

While the aforementioned integration schemes perform well prior to reaching the failure surface, none of properly handle the stress drift away from the failure surface. The reasons for this are twofold. First, hypoplastic models are very sensitive to step size used in stress integration. A too large step size may impair the convergence and stability of the numerical calculation. Second, hypoplastic model allows some stress state outside the failure surface [40]. For some well-defined hypoplastic models, Wu and Niemunis [39] showed that all accessible stress states are within a bound surface. However, the bound surface is often too far

✉ Wei Wu
wei.wu@boku.ac.at

1 Institut für Geotechnik, Universität für Bodenkultur Wien, Feistmantelstrasse 4, 1180 Vienna, Austria

2 Faculty of Engineering, China University of Geosciences, Lumo Road 388, Wuhan 430074, China

from the failure surface, which gives rise to too high strength [40]. Consequently, the error resulted from stress drift away from the failure surface can accumulate in numerical computations and eventually lead to unphysical behaviour [3, 22] and the loss of stability for a boundary value problem. For numerical calculations, it makes sense not to allow stress to wander outside the failure surface. A solution to this problem is to use the return mapping method originally proposed for elastoplastic models [12]. In this work, we will investigate when the return mapping is needed and how the return mapping can improve accuracy. In doing so, we consider a fairly general and simple hypoplastic model so that our approach can be easily adopted to handle more sophisticated hypoplastic models.

In Sect. 2, the hypoplastic constitutive model and its numerical equations are outlined. In Sect. 3, several commonly used integration schemes are introduced. In Sect. 4 the performance of the integration algorithms is examined by performing a series of triaxial compression tests. The significance of the stress correction at the failure surface is shown by two boundary value problems.

2 A simple hypoplastic constitutive model

2.1 Constitutive model

In the framework of hypoplasticity, the constitutive equation is written in two parts, representing, respectively, the reversible and irreversible behaviour. We start with the formulation by Wu and Kolymbas [38] and write the hypoplastic rate-constitutive equation as the sum of the linear and nonlinear terms of the strain rate $\dot{\boldsymbol{\varepsilon}}$

$$\mathring{\boldsymbol{\sigma}} = \mathscr{L}(\boldsymbol{\sigma}) : \dot{\boldsymbol{\varepsilon}} - N(\boldsymbol{\sigma})\|\dot{\boldsymbol{\varepsilon}}\| \tag{1}$$

where the terms $\mathscr{L}$ and N, respectively, denote the linear and nonlinear components, $\boldsymbol{\sigma}$ is the Cauchy stress tensor, and $\dot{\boldsymbol{\varepsilon}}$ is the stretching tensor. $\|\dot{\boldsymbol{\varepsilon}}\| = \sqrt{\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)}$ stands for the Euclidean norm. The Jaumann rate of the Cauchy stress tensor $\mathring{\boldsymbol{\sigma}}$ is defined in terms of the time-derivative of the Cauchy stress tensor $\dot{\boldsymbol{\sigma}}$ and the spin tensor $\boldsymbol{w}$

$$\mathring{\boldsymbol{\sigma}} = \dot{\boldsymbol{\sigma}} + \boldsymbol{\sigma} \cdot \boldsymbol{w} - \boldsymbol{w} \cdot \boldsymbol{\sigma} \tag{2}$$

The stretching and spin tensors are related to the velocity gradient tensor through

$$\dot{\boldsymbol{\varepsilon}} = \frac{1}{2}(\nabla \boldsymbol{v} + \boldsymbol{v}\nabla), \quad \boldsymbol{w} = \frac{1}{2}(\nabla \boldsymbol{v} - \boldsymbol{v}\nabla) \tag{3}$$

where $\boldsymbol{v}$ is the velocity and ∇ is the gradient operator.

Within the above framework, a simple critical state hypoplastic constitutive equation for granular materials is proposed by Wu [42], which is an improvement of an early version by Wu and Bauer [37]. This model consists of

three linear terms and one nonlinear term in the stretching tensor $\dot{\boldsymbol{\varepsilon}}$.

$$\mathring{\boldsymbol{\sigma}} = C_1(\mathrm{tr}\boldsymbol{\sigma})\dot{\boldsymbol{\varepsilon}} + C_2(\mathrm{tr}\dot{\boldsymbol{\varepsilon}})\boldsymbol{\sigma} + C_3 \frac{\mathrm{tr}(\boldsymbol{\sigma} \cdot \dot{\boldsymbol{\varepsilon}})}{\mathrm{tr}\boldsymbol{\sigma}}\boldsymbol{\sigma}$$
$$+ C_4(\boldsymbol{\sigma} + \boldsymbol{\sigma}^*)\|\dot{\boldsymbol{\varepsilon}}\|I_e, \tag{4}$$

in which C_i ($i = 1, 2, 3, 4$) are dimensionless parameters. The deviatoric stress tensor $\boldsymbol{\sigma}^*$ in Eq. (4) is defined by $\boldsymbol{\sigma}^* = \boldsymbol{\sigma} - 1/3(\mathrm{tr}\boldsymbol{\sigma})\boldsymbol{\delta}_{ij}$, with $\boldsymbol{\delta}_{ij}$ being the Kronecker delta. I_e is adopted as the critical state function. It is through I_e that the model captures the effects of density and confining pressure on the strain–stress behaviour. Several forms for I_e can be found in the literature [7, 15, 21, 41], but in the present work, a different formulation for the critical state function I_e is proposed:

$$I_e = \left(\frac{e}{e_{crt}}\right)^\alpha \tag{5}$$

where e and e_{crt} refer to the current void ratio and critical state void ratio, respectively, and α is a constitutive constant that controls the degree of strain softening. The critical state function has the value $I_e = 1$ at the critical state, greater than 1 for a loose state, and less than 1 for a dense state. The evolution of the void ratio follows the evolution of the volumetric strain, $\dot{\varepsilon}_v = \mathrm{tr}(\dot{\boldsymbol{\varepsilon}})$, according to the following relationship:

$$\dot{e} = (1 + e)\dot{\varepsilon}_v; \tag{6}$$

The critical state void ratio is calculated according to Li and Wang [13]. A slightly modified form is used in this work:

$$e_{crt} = e_{co}\exp\left[-\lambda(\frac{p}{p_a})^\xi\right] \tag{7}$$

where e_{co}, λ, and ξ are parametric constants, and $p = \mathrm{tr}\boldsymbol{\sigma}/3$ and p_a denote the hydrostatic pressure and atmospheric pressure (101.325 kPa) for normalisation, respectively. With the additional term I_e, we ensure that both dense and loose sands can be modelled with only one set of material parameters.

Although hypoplastic Eq. (4) was developed mainly for cohesionless soils, in practice, most soils show cohesion to some extent. However, the constitutive model is able to take cohesion into consideration by simply replacing the stress tensor $\boldsymbol{\sigma}$ with the following translated stress tensor [33, 42]:

$$\boldsymbol{\sigma}_c = \boldsymbol{\sigma} - p_t\boldsymbol{\delta}_{ij} \tag{8}$$

where the translated scale $p_t = c/\tan\phi$ and the parameters c and ϕ are, respectively, the cohesion and friction angle of the cohesive soil.

2.2 Explicit form of the failure and bound surfaces

The failure criterion is defined by vanishing stress rate for a non-vanishing strain rate. By assuming that N contains the critical state function I_e, we have:

$$\mathring{\boldsymbol{\sigma}} = (\mathscr{L} - \boldsymbol{N} \otimes \vec{\boldsymbol{\varepsilon}}) : \dot{\boldsymbol{\varepsilon}} = 0 \tag{9}$$

where $\vec{\boldsymbol{\varepsilon}} = \dot{\boldsymbol{\varepsilon}} / \|\dot{\boldsymbol{\varepsilon}}\|$ stands for the direction of strain, and the symbol $\otimes$ denotes an outer product between two tensors. By making use of the fact that $\vec{\boldsymbol{\varepsilon}} : \vec{\boldsymbol{\varepsilon}} = 1$, the failure criterion can be readily derived:

$$f(\boldsymbol{\sigma}) = \|\mathscr{L}^{-1} : \boldsymbol{N}\| - 1 = 0, \tag{10}$$

Therefore, the condition of invertibility is $\|\mathscr{L}^{-1} : \boldsymbol{N}\| < 1$. This condition is identical with the requirement that a stress state $\boldsymbol{\sigma}$ should lie inside the failure surface given in Eq. (10), which means constitutive Eq. (4) is invertible when the stress lies inside the failure surface.

The explicit formula of the failure surface can be obtained using the symbolic computational programme *Mathematica*, which gives rise to the failure surface:

$$f(\boldsymbol{\sigma}) = \sqrt{J_2} + \varsigma I_1 = 0, \tag{11}$$

where J_2 and I_1 are, respectively, the second deviatoric stress invariant and the first stress invariant, and ς is a constant determined by the dimensionless parameters $C_i (i = 1, 2, 3, 4)$ and the critical state function I_e, which can be found in "Appendix".

The hypoplastic model is also characterised by a bound surface that bounds the accessible stress states and which can be derived based on the procedure by Wu [36] and Wu and Niemunis [40]. The explicit formulation of the bound surface is expressed as follows:

$$\mathcal{B}(\boldsymbol{\sigma}) = \sqrt{J_2} - \frac{C_1}{\sqrt{(I_e C_4)^2 - 3C_1^2}} I_1 = 0, \tag{12}$$

With the help of Eqs. (11) and (12), the failure and bound surfaces can be plotted once the model parameters are given.

2.3 Second-order work and stability surface

In view of the complexity of the nonlinear constitutive models, it is desirable to obtain qualitative properties such as stability and uniqueness of the boundary value problems posed with the constitutive models [40]. The problem of stability can be approached based on the analysis of the second-order work. Instability considered in terms of the second-order work virtually means the possibility for spontaneous increase in kinetic energy of the body due to a small disturbance in velocity [19]. According to Hill [10], a sufficient condition for stability is the second-order work

$$\mathcal{W}_2(\boldsymbol{\sigma}) = \text{tr}(\mathring{\boldsymbol{\sigma}}\dot{\boldsymbol{\varepsilon}}) > 0 \tag{13}$$

for all directions of stretching. For hypoplastic models, the second-order work can become negative before the failure surface [35]. Let us consider constitutive equation (4) and search for the boundary between positive and negative second-order work by letting $\mathcal{W}_2 = 0$:

$$\mathcal{W}_2(\boldsymbol{\sigma}) = \dot{\boldsymbol{\varepsilon}} : (\mathscr{L} - \boldsymbol{N} \otimes \vec{\boldsymbol{\varepsilon}}) : \dot{\boldsymbol{\varepsilon}} = 0 \tag{14}$$

If this boundary builds a surface in the stress space, it will be called stability surface. By using the analytical approach proposed by Niemunis [18], an explicit expression of the stability surface can be readily derived to be:

$$\begin{aligned}
\mathcal{W}_2(\boldsymbol{\sigma}) = {} & (L_1^2 - N_1^2)\left[4L_3(L_3 + L_1) + N_1^2\right]^2 \\
& - N_3^2\left[16L_1^4 - 32L_1^3 L_3 + 3N_1^4 + 3N_1^2 N_3^2 + N_3^4 \right. \\
& + 2L_1 L_3(4L_3^2 + 19N_1^2 - 4N_3^2) \\
& \left. - L_3^2(20N_1^2 + N_3^2) + 4L_1^2(2L_3^2 - 5N_1^2 + 2N_3^2)\right]
\end{aligned} \tag{15}$$

where the variables are expressed as function of the stress invariants I_1 and J_2: $L_1 = C_1 I_1$, $L_3 = (C_1 + C_2 + C_3/3) I_1 + 2C_3 J_2 / I_1$, $N_1 = -C_4 I_1 \sqrt{2J_2} / \sqrt{I_1^2 + 6J_2}$, and $N_3 = C_4(I_1^2 + 12J_2) / \sqrt{3I_1^2 + 18J_2}$. Note that the stability surface depends on the density through the critical state function I_e by replacing N_1 and N_3 with $I_e N_1$ and $I_e N_3$, respectively. The implication of the stability surface for FE calculations is beyond the scope of this paper. In general, the stability should be considered within a BVP rather than an element. Some stability indicators can be introduced to avoid pitfalls in the FE calculations [23].

By using Eq. (15), we plot the stability surface together with the bound and failure surfaces in the principal stress space. As shown in Fig. 1, the stability surface of the hypoplastic model is a cone with its apex at the origin in the principal stress space. The bound and failure surfaces possess similar geometry, outside the stability surface. Wu and Niemunis [40] show that a stress state may lie outside the failure surface for some paths. We proceed to investigate how far a stress state can drift away from the failure surface for constitutive equation (4).

Let us consider three stress states $\boldsymbol{\sigma}_b$, $\boldsymbol{\sigma}_f$ and $\boldsymbol{\sigma}_s$ with the same Lode angle lying on the bound, the failure, and the stability surfaces, respectively. The corresponding stress ratio can be obtained as follows:

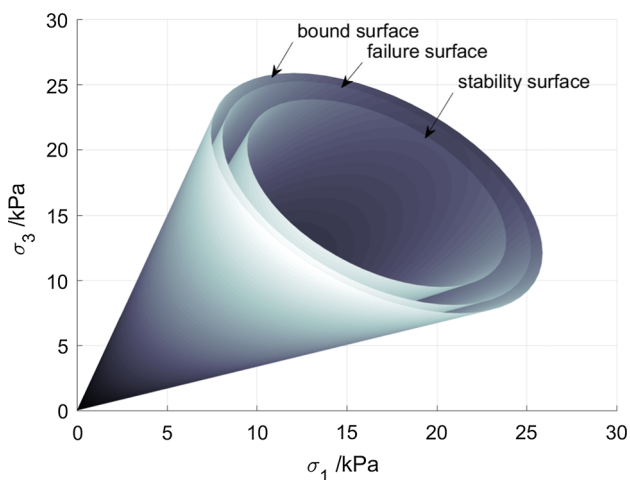$$\frac{\sigma_1 - \sigma_3}{\sigma_1 + \sigma_3} = \sin\phi_{\text{mob}} \tag{16}$$

Fig. 1 The bound, failure and stability surfaces in principal stress space

Table 1 The mobilised friction angles ϕ_{mob} at different stress states

Stress state	Triaxial compression[/°]			Triaxial tension[/°]		
ϕ_{mob_b}	21.4	28	35.9	28.8	49.2	73.2
ϕ_{mob_f}	20.0	25.0	30.0	26.3	36.8	43.6
ϕ_{mob_s}	11.7	16.8	21.9	13.5	21.0	29.8

step [40]. However, a stress state may also lie between the failure and bound surface for some strain paths. Table 1 shows the mobilised friction angles for the failure surface and bound surface in triaxial tests. We observe that the difference of the mobilised friction angles between the failure and bound surface increases with the critical friction angle and may reach 30° for triaxial extension, which is far too high to be accepted. Obviously, stress corrections are needed to bring the stress back to the failure surface.

3 Time integration of the constitutive model

3.1 Tangential stiffness for FEM

The finite element formulation for the linear continuum model follows the standard Galerkin approximation of the weak form based on the virtual work principle. Thus the global finite element equation can be expressed:

$$\sum \int_{V_e} \mathrm{B}^{\mathrm{T}} \boldsymbol{\sigma}(\boldsymbol{u}_n) dV_e = \mathrm{P}_{n\_\mathrm{ext}} \tag{17}$$

where n denotes the nth increment of the nonlinear analysis. $\boldsymbol{P}_{n\_\mathrm{ext}}$ is the nodal force vector, and $\boldsymbol{u}$ is the unknown nodal displacement vector. Thus the unbalanced force vector can be expressed as:

$$\mathrm{F}(\boldsymbol{u}) = \sum \int_{V_e} \mathrm{B}^{\mathrm{T}} \boldsymbol{\sigma}(\boldsymbol{u}_n) dV_e - \mathrm{P}_{n\_\mathrm{ext}} \tag{18}$$

where σ_1 and σ_3 are the maximum and minimum principal stresses, respectively. The stress ratio can be obtained from the mobilised friction angle ϕ_{mob}. In contrast to the Mohr–Coulomb yield criterion, the mobilised friction angle ϕ_{mob} varies for different Lode angles for the hypoplastic model. Obviously, the mobilised friction angles obtained from $\boldsymbol{\sigma}_b$, $\boldsymbol{\sigma}_f$ and $\boldsymbol{\sigma}_s$ are dependent on the Lode angle. Figure 2 shows the three mobilised friction angles ϕ_{mob_b}, ϕ_{mob_f} and ϕ_{mob_s} for different critical friction angles. For the stress state in triaxial compression (the Lode angle is zero) with a 20° frictional angle, one obtains $\phi_{\mathrm{mob}\_b} = 21.4°$, $\phi_{\mathrm{mob}\_f} = 20°$ and $\phi_{\mathrm{mob}\_s} = 11.7°$ while for the stress state in extension (the Lode angle is 60°) the corresponding friction angles are $\phi_{\mathrm{mob}\_b} = 28.8°$, $\phi_{\mathrm{mob}\_f} = 26.3°$ and $\phi_{\mathrm{mob}\_s} = 16.8°$.

The bound surface is an intrinsic property of hypoplastic constitutive equations, which thus some advantage in the numerical integration over most conventional constitutive models, since the stress states lying outside the bound surface will be automatically corrected in the next time

Fig. 2 The mobilised fiction angle ϕ_{mob} for different critical frictional angles ϕ: **a** $\phi = 20°$, **b** $\phi = 25°$, and **c** $\phi = 30°$

For the ith iteration in the framework of the standard Newton–Raphson procedure, the global stiffness matrix reads:

$$\mathrm{K}(\boldsymbol{u}_n^i) = \frac{\partial \mathrm{F}(\boldsymbol{u}_n^i)}{\partial \boldsymbol{u}} = \sum \int_{V_e} \mathrm{B}^\mathrm{T} \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} \Big|_{\boldsymbol{u}_n^i} \mathrm{B} dV_e \qquad (19)$$

For the hypoplastic constitutive model, the Jacobian matrix can be readily written out:

$$\mathrm{D} = \frac{\partial \Delta \boldsymbol{\sigma}}{\partial \Delta \boldsymbol{\varepsilon}} = \mathscr{L} + I_e \boldsymbol{N} \vec{\boldsymbol{\varepsilon}} \qquad (20)$$

For constitutive model (4), the tangential stiffness tensor at a given stress σ_{ij} is given below:

$$\frac{\partial \mathring{\sigma}_{ij}}{\partial \dot{\varepsilon}_{mn}} = C_1 \Theta\, I_{ijmn} + C_2 \sigma_{ij} \delta_{mn} + C_3 \frac{\sigma_{ij} \sigma_{mn}}{\mathrm{tr}\boldsymbol{\sigma}} + I_e C_4 (\sigma_{ij} + \sigma_{ij}^*) \vec{\varepsilon}_{mn} \qquad (21)$$

where I_{ijmn} is the fourth rank identity tensor with components $I_{ijmn} = (\delta_{im}\delta_{jn} + \delta_{jm}\delta_{in})/2$. It is worth noting that the Jacobian matrix may become negative definite when using the above tangential operator. A numerical study on the performance of different tangential operators can be found in [4].

3.2 Stress integration algorithms

The constitutive equation can be regarded as an ordinary differential equation, for which the general time integration over an increment step $t \in [t_n, t_{n+1}]$ can be written as:

$$\boldsymbol{\sigma}_{n+1} = \boldsymbol{\sigma}_n + \int_{t_n}^{t_{n+1}} h(\boldsymbol{\sigma}, e, \dot{\boldsymbol{\varepsilon}}) dt = \boldsymbol{\sigma}_n + \mathrm{D}\Delta\boldsymbol{\varepsilon}, \quad n = 1, 2, \ldots \qquad (22)$$

According to Eq. (6), a closed form of integration for the void ratio is available:

$$e_{n+1} = (1 + e_n) \cdot \exp(\Delta \dot{\varepsilon}_v) - 1 \qquad (23)$$

The solution for Eqs. (22) and (23) can be obtained stepwise according to one of the integration algorithms in the next subsection.

3.2.1 Generalised midpoint algorithms

Following the method proposed in [5, 6], all stress components and state variables are collected in vector $\boldsymbol{y}$ for the simultaneous integration of Eqs. (22) and (23):

$$\boldsymbol{y} = \{\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}, \sigma_{23}, g_1 \ldots g_m\}^\mathrm{T} \qquad (24)$$

where $g_i (i = 1 \ldots m)$ are additional state variables, e.g. void ratio. The integration of Eq. (24) requires the solution of initial value problem:

$$\boldsymbol{y}'(t) = H(\boldsymbol{y}(t)), \quad \boldsymbol{y}'(0) = \boldsymbol{y}(0). \qquad (25)$$

Generalised midpoint algorithms are among the most widely used second-order integration methods [2]. The general form of the generalised midpoint method (e.g. [30]) can be written as:

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \Delta t_{n+1}[(1 - \theta)\dot{\boldsymbol{y}}_n + \theta \ddot{\boldsymbol{y}}_{n+1}] \quad n = 1, 2, \ldots \qquad (26)$$

where $\Delta t_{n+1} = t_{n+1} - t_n$ is the time step increment and the parameter θ is a prescribed constant within [0, 1]. Note that generalised midpoint algorithms with values of θ equal to 1 and 0, respectively, correspond to the implicit backward Euler method and explicit forward Euler method with the Crank–Nicolson (midpoint or trapezoidal) method obtained by setting $\theta = 0.5$. Our implementation tests three widely used schemes: the explicit forward Euler method (FE), the modified Euler method (ME) and the implicit Crank–Nicolson method (CN).

3.2.2 Adaptive explicit algorithms

The accuracy of an integration method can be improved by reducing the size of the time increment. Although this can be carried out in a straightforward manner by dividing the time increment into several equal substeps, the better accuracy is usually gained at the cost of computational time or failure of error control within a tolerant range. A more powerful approach is to employ the adaptive integration schemes described in the literature [27, 34], which enable users to adjust the substep size automatically according to the local truncation error. Studies have revealed that this approach has the merits of being efficient and robust for a wide range of constitutive models. In the present paper, several adaptive explicit integration methods, namely the modified Euler substepping scheme (MEsec), the Richardson extrapolation substepping scheme (REsec) and the Runge–Kutta–Fehlberg substepping scheme (RKFsec), are implemented and compared.

To compute the local error at each substep of the stress integration, two approximate solutions with different orders of accuracy (p, q) are obtained and compared. If the two solutions are in close agreement, the approximation is accepted; otherwise, the substep is rejected and the corresponding step size is further reduced. For the generic substep k in the time interval $[t_n, t_{n+1}]$, with dimensionless size $\Delta T_k \in (0, 1)$ given by the following equation:

$$\Delta T_k = (t_{k+1} - t_k)/(t_{n+1} - t_n) \leqslant 1 \quad \text{and} \quad \sum_{k=1}^{n_s} \Delta T_k = 1, \qquad (27)$$

two different approximate solutions of the evolution problem (25) are obtained simultaneously according to

$$\boldsymbol{y}_{k+1}^{(p)} = \boldsymbol{y}_k + \Phi_1(\boldsymbol{y}_k, \Delta\boldsymbol{\varepsilon}_{n+1}, \Delta T)\Delta(T_k)^{(p+1)} \qquad (28a)$$

$$\boldsymbol{y}_{k+1}^{(q)} = \boldsymbol{y}_k + \Phi_2(\boldsymbol{y}_k, \Delta\boldsymbol{\varepsilon}_{n+1}, \Delta T)\Delta(T_k)^{(q+1)} \qquad (28\text{b})$$

The two function Φ_1 and Φ_2 are constructed as follows according to Table 2, which summarises different integration methods:

$$\Phi_1 := \sum_{i=0}^{p} k_i^{(p)} h_i(\boldsymbol{y}_k, \Delta\boldsymbol{\varepsilon}_{n+1}, \Delta T) \qquad (29\text{a})$$

$$\Phi_2 := \sum_{i=0}^{q} k_i^{(q)} h_i(\boldsymbol{y}_k, \Delta\boldsymbol{\varepsilon}_{n+1}, \Delta T) \qquad (29\text{b})$$

where

$$h_i(\boldsymbol{y}_k, \Delta\boldsymbol{\varepsilon}_{n+1}, \Delta T) := h\left(\boldsymbol{y}_k + \Delta T \sum_{j=0}^{l-1} \vartheta_{lj} h_j, \Delta\boldsymbol{\varepsilon}_{n+1}\right) \qquad (30)$$

For the sake of simplicity, we usually set $q = p + 1$. The constants $k_i^{(p)}$, $k_i^{(q)}$ and ϑ_{lj} are used to obtain the approximated solutions of order p and q, respectively. Then the local truncation error of the lower-order method at time T_{k+1} can be obtained by using the difference in the above two approximate solutions:

$$\boldsymbol{R}_{k+1} = \boldsymbol{y}_{k+1}^{(p)} - \boldsymbol{y}_{k+1}^{(q)}, \quad \text{and} \quad R_{k+1} = \frac{\|\boldsymbol{R}_{k+1}\|}{\|\boldsymbol{y}_{k+1}^{(q)}\|} \qquad (31)$$

The integration over the k-th substep is assumed to be successful when, for a given stress error tolerance STOL,

$$R_{k+1} \leqslant \text{STOL}, \qquad (32)$$

with the new substep size then estimated using the following extrapolation formula:

$$\Delta T_{k+1} = \Delta T_k \left[\frac{\text{STOL}}{R_{k+1}}\right]^{1/(p+1)} \qquad (33)$$

If the estimated error is less than the prescribed accuracy tolerance STOL, the step is accepted and we enlarge our step size according to

$$\Delta T_{k+1} = \Delta T_k \cdot \min\left\{1.1, 0.9\left[\frac{\text{STOL}}{R_{k+1}}\right]^{1/(p+1)}\right\} \qquad (34)$$

If condition (32) is not satisfied, the k-th substep will be re-calculated with a smaller step size ΔT_k^*:

$$\Delta T_k^* = \Delta T_k \cdot \max\left\{0.25, 0.9\left[\frac{\text{STOL}}{R_{k+1}}\right]^{1/(p+1)}\right\} \qquad (35)$$

Note that the right side of Eq. (33) is multiplied by a factor that is typically set to 0.9. An upper bound of 1.1 and a lower bound of 0.25 are also taken for each new substep in order that the extrapolation is not carried too far. More details regarding the substepping algorithm can be found in the literature [1, 29, 30]. After the integration process is complete, the stress tensor can be extracted from the vector $\boldsymbol{y}$.

3.2.3 Correction of stresses to failure surface

At the end of each increment in the integration process, the stresses may diverge from the failure function so that $f(\boldsymbol{\sigma}) > FTOL$. The extent of this violation, which is commonly known as failure surface 'drift', depends on the accuracy of the integration scheme and the nonlinearity of the constitutive relations. Sloan [27] suggested that, provided the integration is performed accurately, the extent of drift from the failure surface tends to be small and no remedial action is required. Wu and Niemunis [40] and Niemunis [18], on the other hand, have reported that some stress states may surpass the failure surface irrespective of the accuracy of the used integration method. In such cases, the stress state does not satisfy the consistent condition.

Let us consider a stress state inside the failure surface, e.g. stress $\boldsymbol{\sigma}_n$ at the nth step. As demonstrated in [39, 40], the hypoplastic model allows some stress state lying outside the failure surface. For the strain path shown in Fig. 3,

Table 2 Different integration methods for generating estimated errors

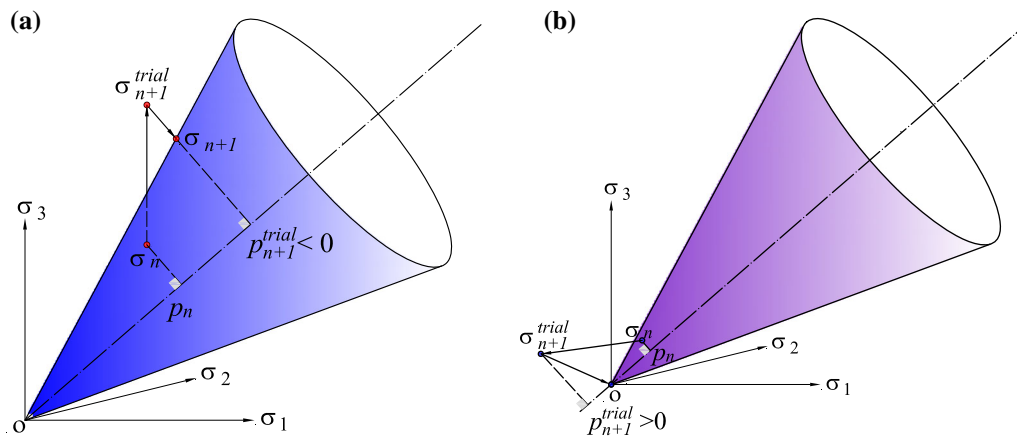| Method	Formulation	Error estimation
FE	$y_{n+1} = y_n + \Delta y_{n+1}$	$R_{n+1} = \frac{\|y_{n+1} - y_n\|}{y_{n+1}}$
ME	$y_{n+1}^1 = y_n + \Delta y_{n+1}^1$ $y_{n+1}^2 = y_n + 0.5(\Delta y_{n+1}^1 + \Delta y_{n+1}^2)$	$R_{n+1} = \frac{\|y_{n+1}^2 - y_n^1\|}{y_{n+1}^2}$
CN	$y_{n+1} = y_n + 0.5(\Delta y_n + \Delta y_{n+1})$	$R_{n+1} = \frac{\|y_{n+1} - y_n\|}{y_{n+1}}$
RE	$y_{n+1}^1 = y_n + \Delta y_{n+1}^1$ $y_{n+1}^2 = y_n + 0.5\Delta y_{n+1}^1$ $y_{n+1}^3 = y_n^2 + 0.5\Delta t_{n+1} h(y_{n+1}^2)$	$R_{n+1} = \frac{\|y_{n+1}^3 - y_n^1\|}{y_{n+1}^3}$
RKF$_{23}$	$y_{n+1}^1 = y_n + 0.5\Delta y_{n+1}^1$ $y_{n+1}^2 = y_n - \Delta y_{n+1}^1 + 2\Delta y_{n+1}^2$ $y_{n+1}^3 = y_n + \Delta y_{n+1}^2$ $y_{n+1}^k = y_n + \frac{1}{6}\Delta y_{n+1}^1 + \frac{2}{3}\Delta y_{n+1}^2 + \frac{1}{6}\Delta y_{n+1}^3$	$R_{n+1} = \frac{\|y_{n+1}^k - y_n^3\|}{y_{n+1}^k}$

Fig. 3 Sketch of return mapping scheme, **a** return to the cone, and **b** return to the apex, the direct of $\boldsymbol{\sigma}$ is negative

no matter how accurate the integration method, the stress defined by $\boldsymbol{\sigma}_{n+1}^{\text{trial}}$ at the $(n+1)$th step of analysis will violate the consistent condition, so that

$$f(\boldsymbol{\sigma}_{n+1}^{\text{trial}}) = \sqrt{J_2(\boldsymbol{\sigma}_{n+1}^{*\text{trial}})} + \varsigma(C_i, I_e)I_1 > FTOL, \tag{36}$$

with the explicit formulation defined in "Appendix".

As any deviations from the failure surface are cumulative and may result in unacceptable errors in subsequent computations, the stresses should be corrected to satisfy the current consistent condition. As shown in Fig. 3a, for $p_{n+1}^{\text{trial}} < 0$, the stress is corrected along the radial direction to the failure surface [35]. With the radial return scheme, the corrected stress state takes the following form:

$$\boldsymbol{\sigma}_{n+1}^* = \eta\boldsymbol{\sigma}_{n+1}^{*\text{trial}}, \quad p_{n+1} = p_{n+1}^{\text{trial}}, \tag{37}$$

where η is an unknown multiplier. Using the previous definition of J_2, it follows that:

$$\frac{\boldsymbol{\sigma}_{n+1}^*}{J_2(\boldsymbol{\sigma}_{n+1}^*)} = \frac{\boldsymbol{\sigma}_{n+1}^{*\text{trial}}}{J_2(\boldsymbol{\sigma}_{n+1}^{*\text{trial}})}. \tag{38}$$

In order to return the stress state to the failure surface, it is desirable that the total strain increment, $\Delta\varepsilon$, remains unchanged, since this is consistent with the displacement-based finite element procedure. The corrected stress state in Eq. (37) satisfies the consistency condition. Using Eqs. (36) and (38), together with the assumption that departures from the failure surface are sufficiently small that only one return step is required, the consistent condition is expressed as follows:

$$f = \eta\sqrt{J_2(\boldsymbol{\sigma}_{n+1}^{*\text{trial}})} + \varsigma(C_i, I_e)I_1 = 0, \tag{39}$$

which yields the unknown multiplier η.

On the other hand, if the mean stress $p_{n+1}^{\text{trial}} > 0$, as shown in Fig. 3b, the stress is corrected to the apex of the failure surface

$$\boldsymbol{\sigma}_{n+1} = 0, \quad p_{n+1} = 0, \tag{40}$$

After solving the above equation, we correct the violated stress either to the cone or to the apex of the failure surface via Eqs. (37) and (40).

The stress correction scheme can also be easily adapted to incorporate the effect of critical state and cohesion. To this end, the constant ς including the critical state function I_e can be found in "Appendix". Figure 4 shows the corrected Stress Response Envelope (SRE). A detailed study regarding SRE can be found in [30]. For any stress lying between the failure and bound surfaces, the updated stress is forced back to the failure surface and thus the function of the bound surface is abandoned.

4 Numerical tests for different integration strategies

To obtain an overall assessment of the integration methods presented in Sect. 3, a set of numerical tests is conducted for constitutive equation (4). Firstly, drained and undrained
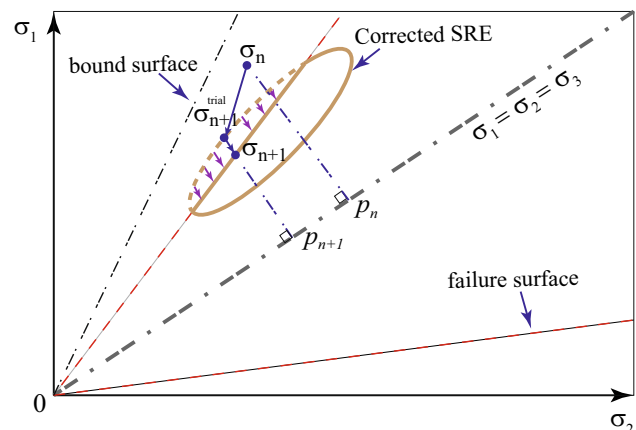

Fig. 4 The sketch of corrected Stress Response Envelope (SRE)

triaxial compression tests are modelled. We study the influence of stress correction on the stress–strain relation in drained and undrained triaxial tests, together with an analysis of the stress drift from the failure surface. Secondly, two boundary problems, namely a rigid footing, and the safety factor of a homogeneous slope, are considered using the finite element code Abaqus Standard with user-defined materials (Umat). The performance of different integration methods and stress correction scheme in solving these problems is then compared in terms of accuracy, efficiency, and robustness.

4.1 Performance of integration methods

Two groups of integration methods are analysed in the numerical triaxial compression tests:

(1) Constant substep method: the forward Euler method (FE), the modified Euler method (ME), and the Crank–Nicolson method (CN);

(2) Adaptive explicit method: the adaptive modified Euler method (MEsec), the adaptive Richardson extrapolation method (REsec), and the adaptive Runge–Kutta–Fehlberg method (RKF23sec).

The formulation of different integration methods and error estimations are shown in Table 2.

To assess their numerical performance, a benchmark solution is obtained using the adaptive RKF45 method, in which the integration error tolerance is set to 10^{-9}. The relative error is calculated for every step as follows:

$$R_n = \frac{\|\boldsymbol{y}_{\text{exact}}^n - \boldsymbol{y}^n\|}{\|\boldsymbol{y}_{\text{exact}}^n\|}. \qquad n = n\text{th step} \tag{41}$$

For the triaxial compression tests, an initial isotropic stress state with $\sigma_{11} = \sigma_{22} = \sigma_{33} = 100$ kPa is assumed. The initial void ratio is set to $e_i = 0.78$ for the drained triaxial test and $e_i = 0.93$ for the undrained triaxial test; both tests are strain-controlled with a maximum axial (vertical) strain of 10% applied, and the horizontal strain increment is calculated by the constitutive model for a given axial strain increment. The parameters used in the simulations are shown in Table 3. In the numerical procedures, two types of increments are adopted. In the first calculation, the loading process is divided into 10 equal increments, representing a large increment size scheme. In the second calculation, the loading process is divided into 20 equal

increments, representing a relatively small increment size scheme. For each scheme, different substeps are performed for the explicit Euler method and the implicit CN method, and different STOLs applied for the CN method. Similarly, the integration error tolerance STOL is changed for the adaptive explicit methods. For each method, the integration results are accepted once convergence is obtained or the iteration number limit reached. The numerical results of the drained and undrained triaxial tests carried out using the various integration methods with 10 increments (2 substeps) and 20 increments (1 substeps) are shown in Fig. 5. The maximum errors obtained from computations with 10 increments under different substeps (substep = 2, 20, 100) for constant substep methods and under different STOLs for adaptive explicit methods are shown in Fig. 6.

As shown in Figs. 5 and 6, the integration strategies exhibit very different behaviour. The forward Euler (FE) method with 2 substeps provides the roughest estimation of the stress–strain response in both the drained and undrained triaxial tests. Indeed, the relative error produced by this scheme reaches 0.029 and 0.6491 in the drained and undrained tests, respectively, which can easily lead to unacceptable results in finite element calculations. The poor performance in the undrained tests is due to the used hypoplastic constitutive model which is very sensitive to both time step and stress path. Comparison of Figs. 5b and 4d reveals that the main error results from the first increment, and this error is accumulated to the rest increments. However, by increasing the number of substeps, the maximum error resulted from FE method can be largely decreased, as shown in Fig. 6a.

The performance of the simple Modified Euler method (ME) is rather poor in undrained triaxial test. The implicit Crank–Nicolson method (CN) with 2 substeps is more accurate than either of the above explicit methods. Among the three adaptive explicit methods, the modified Euler method and the RKF23 method achieve the highest accuracy, with relative errors around of 10^{-5}, whereas the accuracy of the REsec method with STOL = 10^{-1} is relatively low. When STOL is decreased, all adaptive explicit methods perform well; the stress errors decrease as STOL is reduced, with the best accuracy achieved with the error tolerance set to 10^{-6}, as shown in Fig. 6b. In all, the performance of the integration methods is better in the drained tests than in the undrained tests, which implies that the integration methods are dependent on the stress path.

Table 3 Parameters for numerical simulation of the drained triaxial test

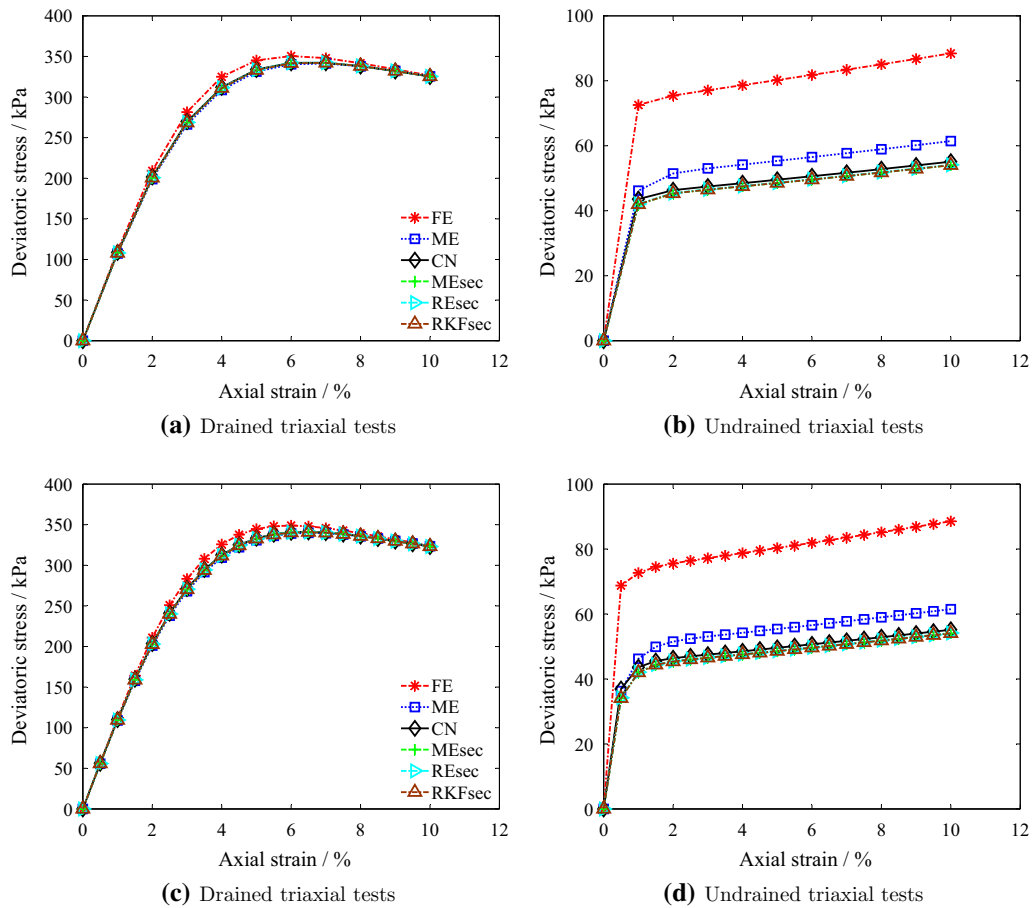| Parameter	C_1	C_2	C_3	C_4	e_{co}	λ	ξ	α
Value	-30.56	-97.11	-286.46	-93.56	0.957	0.022	0.061	1.2

Fig. 5 Stress–strain relations of triaixal compression tests with **a**, **b** 10 increments, 2 substeps, STOL $= 10^{-4}$ and **c**, **d** 20 increments, 1 substeps, STOL $= 10^{-4}$
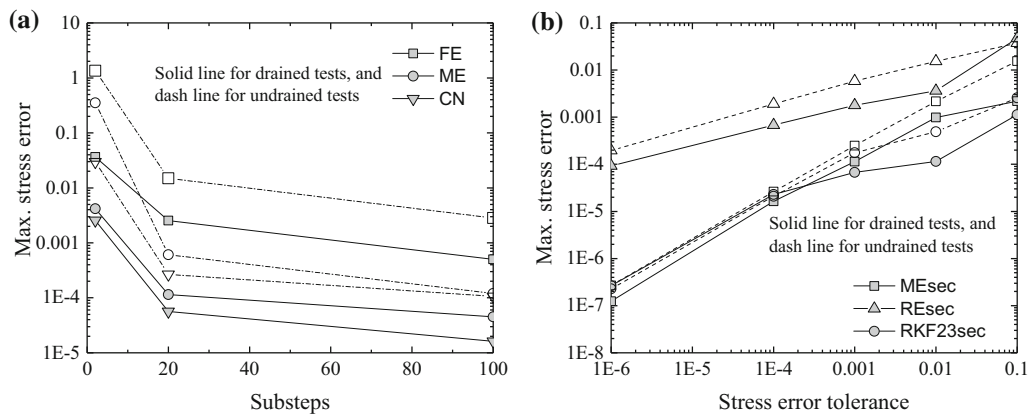

Fig. 6 The maximum stress error obtained from computaion with 10 increments using **a** Constant substep methods, and **b** Adaptive explicit methods

4.2 The effect of stress correction

To evaluate the effect of stress correction, three different initial isotropic stress states with $\sigma_{11} = \sigma_{22} = \sigma_{33} = 100/200/300$ kPa are assumed. The initial void ratio is set to e_i

$= 0.78$ for the drained triaxial test and $e_i = 0.95$ for the undrained triaxial test, with both tests strain-controlled with a maximum axial (vertical) strain of 20%. The parameters used in these simulations are shown in Table 3. We consider only the adaptive RKF23sec method. The

integration error tolerance is set at STOL = 10^{-4}. Stress correction is employed in each step to bring the stress back to the failure surface.

Figure 7 presents the results of stress correction in both the drained and undrained triaxial tests. As can be observed from Fig. 7a, stress correction takes effect only after the peak in the stress–strain curve, which implies that stress correction is particularly relevant in the softening regime. However, stress correction does not make noticeable difference in the undrained test, as shown in Fig. 7b. The stress path recorded in the drained and undrained tests are presented in Fig. 8a. It can be observed from Fig. 8a that whereas the undrained stress paths do not surpass the critical state line, the drained stress paths not only exceed the critical state line but also reach the peak stress state line. This reveals that stress correction mainly takes place in the domain between the critical state line and the peak stress state line. Figure 8b shows that the failure function is violated, i.e. $f > 0$ beyond the peak. However, with the adoption of the stress correction scheme, the failure function f becomes null, thus guaranteeing that the stress lies on the failure surface.

As discussed in Sect. 2, stress drift is particularly relevant for large critical friction angle, e.g. $\phi = 30°$. In order to explore this phenomenon, the effect of stress correction for materials with various friction angles is thus investigated. Figure 9a, which illustrates the stress–strain relationship for different friction angles, reveals that the magnitude of the corrected stress increases with an increase in the friction angle. Figure 9b shows the relative error of stress correction in the drained triaxial tests. The relative error increases from 0.5% for $\phi = 15°$ to 4.5% for $\phi = 45°$. Obviously, stress correction is important in the numerical implementation of the hypoplastic constitutive models.

4.3 Boundary value problems

We now consider two typical boundary value problems, i.e. rigid footing and slope stability. In the rigid footing test, attention is focused not only on accuracy and robustness, but also on the computational efficiency of the numerical schemes. The second problem of slope stability is particularly relevant to bring out the effect of stress correction for the numerical calculations. As shown in the last section,
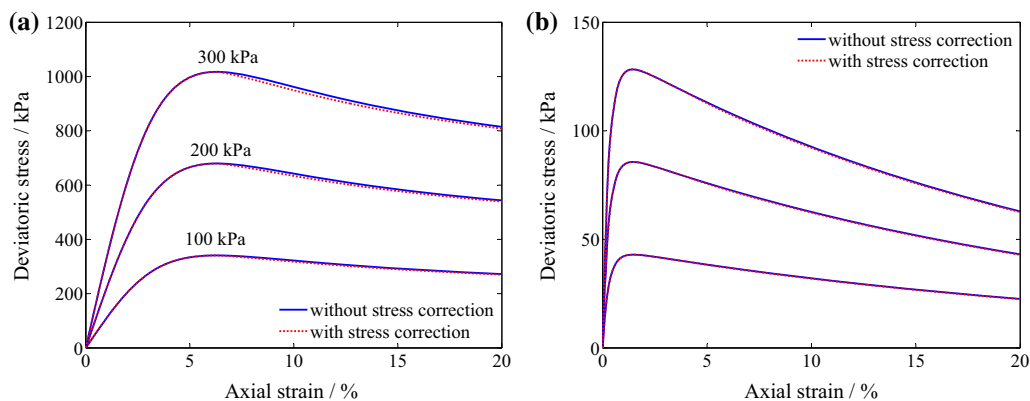

Fig. 7 Results of stress correction in triaxial tests: Stress–strain relations of **a** drained triaxial test, **b** undrained triaxial test
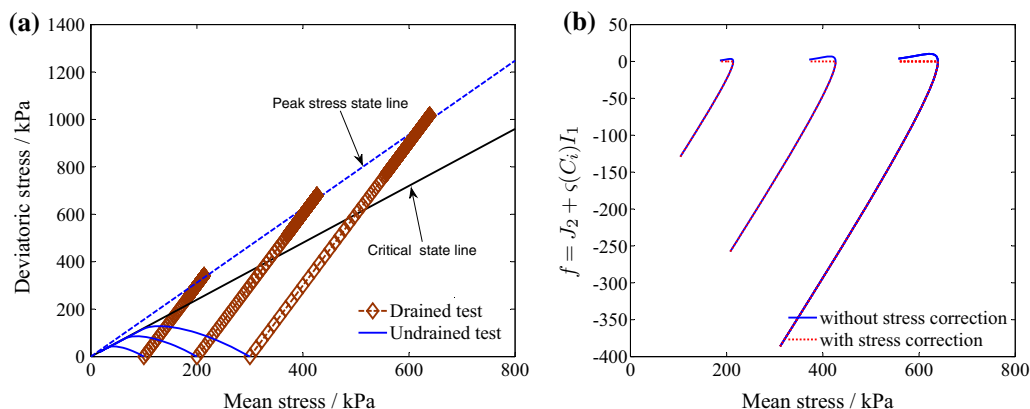

Fig. 8 **a** Stress path of triaxial tests, **b** the value of failure function f in drained triaxial test
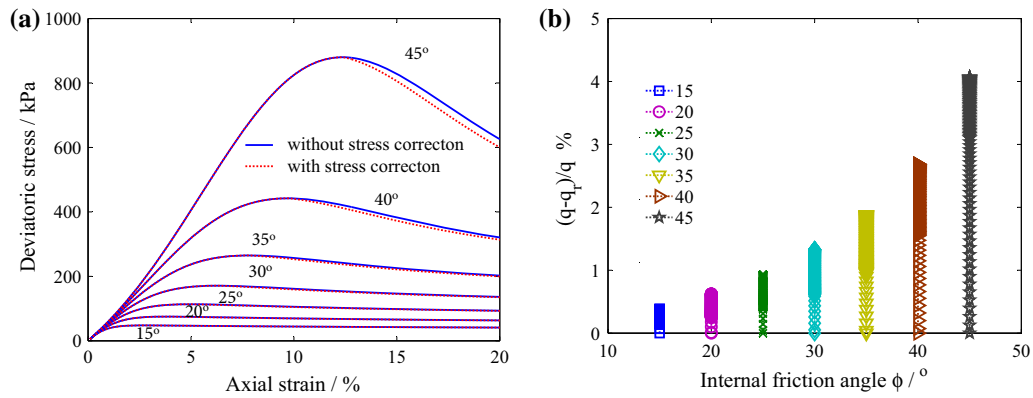
Fig. 9 The effect of stress correction on drained triaxial tests ($p = 50$ kPa): **a** Stress–strain relation with various internal friction angles, and **b** relative error after stress correction

the simple explicit and implicit methods with large step sizes can produce large errors, a constant substepping method is considered together with the forward Euler method, modified Euler method and Crank–Nicolson method. For the latter, different stress tolerances are used and the maximum number of local iterations is set to 10,000. For the adaptive explicit methods, the maximum number of substeps is less than 10,000 and the minimum substeps size is less than 10^{-7} of the current increment size. For both problems, benchmark solutions are obtained via the RKF45sec method with STOL = 10^{-9}, which is compared to the numerical solutions obtained by the different methods with and without stress correction. To accommodate tensile stresses, a cohesion c is assigned to the soil, thereby allowing the development of tensile stresses during the computations.

4.3.1 Rigid footing test

Further investigation of the above numerical methods is carried out for the boundary value problem of a rigid footing. The computation domain, as shown in Fig. 10, is 4.0 m deep by 12 m wide and the width of the footing is $w = 1.2$ m.

For the sake of simplicity, an asymmetric model is chosen using a total number of 150 four-node plane strain elements, and 600 Gauss integration points. The maximum vertical displacement is $d = 0.5$ m, at which point the vertical force reaches its peak value, with the displacement divided into 100 equal increments. Prior to loading of the footing, an initial geostatic stress (120 kPa) is applied. The parameters used in this simulation are shown in Table 4 for an initial void ratio of $e_i = 0.78$. Stress integration errors are evaluated at the end of calculation. Note that the explicit Euler method with stress correction is not involved in the evaluation of stress error. The numerical results displayed in Table 5, in which the "Total number of
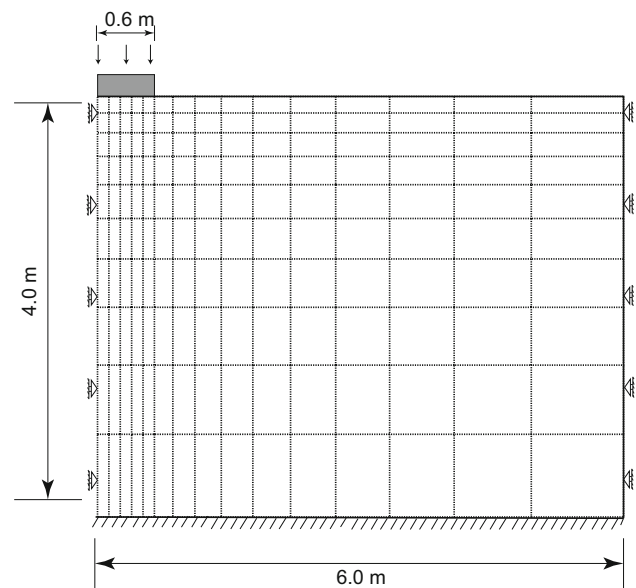

Fig. 10 Finite element meshes of the rigid footing

substeps" is calculated according to the accumulated number of substeps for all Gauss points across all increments, while the "Maximum number of substeps" represents the substeps of one Gauss point with the maximum number of substeps across all increments.

Table 5 shows that while the FE method and ME method have similar average errors, the latter method requires twice the CPU time of the former. In addition, the stress correction accounts for a very small proportion of the total CPU time. In contrast, the CN method with 20 substeps produce less accurate result than the CN method with 100 substeps, although the former took less time than the latter. As expected, the adaptive explicit methods are able to control the integration error and CPU time cost effectively for a given STOL. Among the two adaptive methods, the MEsec method is more efficient and the RKF23sec method more accurate, both with excellent performance.

Table 4 Parameters for the rigid footing test

Parameter	C_1	C_2	C_3	C_4	e_{co}	λ	ξ	α	c (kPa)
Value	− 50.1	− 520.7	− 1802.3	− 300.6	0.957	0.122	0.061	1.5	41.6

Table 5 Results of different methods for rigid footing test (100 increments)

Method	CPU time (s)	Total number of substeps	Maximum number of substeps	Average error
FE (100 substeps)	22.5	6×10^6	10,000	1.928×10^{-4}
FEs (100 substeps)	23.7	6×10^6	10,000	–
ME (100 substeps)	41.8	6×10^6	10,000	1.293×10^{-4}
MEs (100 substeps)	43.1	6×10^6	10,000	
CN method (20 substeps)				
STOL = 10^{-1}	16.8	1.2×10^6	2000	2.886×10^{-4}
STOL = 10^{-2}	15.9	1.2×10^6	2000	2.886×10^{-4}
STOL = 10^{-3}	22	1.2×10^6	2000	2.886×10^{-4}
STOL = 10^{-4}	17.2	1.2×10^6	2000	2.886×10^{-4}
CN method (100 substeps)				
STOL = 10^{-1}	38	6×10^6	10,000	1.928×10^{-4}
STOL = 10^{-2}	43.5	6×10^6	10,000	1.928×10^{-4}
STOL = 10^{-3}	41.9	6×10^6	10,000	1.928×10^{-4}
STOL = 10^{-4}	45.7	6×10^6	10,000	1.928×10^{-4}
MEsec method				
STOL = 10^{-1}	16.6	317,357	1407	1.525×10^{-4}
STOL = 10^{-2}	16.3	647,377	1514	1.983×10^{-5}
STOL = 10^{-3}	16.9	678,211	1486	7.647×10^{-6}
STOL = 10^{-4}	18.2	754,393	1745	2.884×10^{-6}
RKF23sec method				
STOL = 10^{-1}	24.1	84,150	1223	4.076×10^{-5}
STOL = 10^{-2}	29.1	251,326	1351	7.550×10^{-6}
STOL = 10^{-3}	29.5	385,152	1387	1.252×10^{-6}
STOL = 10^{-4}	30.8	627,134	1475	1.914×10^{-7}
RKF45sec method				
STOL = 10^{-9}	29.5	644,161	1429	–

A colour contour plot of the number of substeps for each element is shown in Fig. 11a, which displays a well-defined shear band developed near the footing. As expected, the average number of substeps is much higher in the shear band, which indicates that the substepping scheme reduces the calculation effort. Figure 11b shows a triangular zone of large strain directly under the foundation, as well as a radial zone and a Rankine passive zone, forming together the three zones assumed by Terzaghi.

The relationship between vertical force and vertical displacement is shown in Fig. 12. First, let us look at the calculations with one substep. The calculation without stress correction becomes instable near the ultimate load, while the calculation with stress correction remains stable even beyond the peak. The calculations with 10 constant substeps show that the performance can be improved by substepping, and the stress correction gives rise to slightly lower limit load.

4.3.2 Stability of homogeneous slope

The stress correction scheme is further validated by evaluating the safety factor of a homogeneous slope and simulating the subsequent failure process. In slope stability analysis, the safety factor is typically evaluated using the so-called shear strength reduction technique, in which the shear strength (friction angle and cohesion) is reduced by a reduction factor until slope failure occurs; the safety factor is thus defined by this reduction factor. Here the safety factor and failure of a homogeneous slope are studied using different integration methods, i.e. the implicit CN method, the FE method, the adaptive RKF23 method and each of these
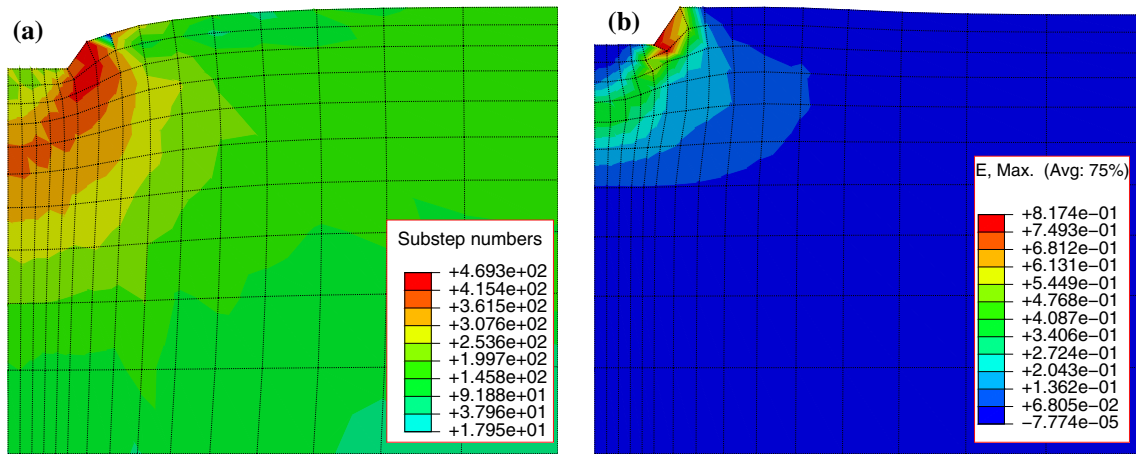
Fig. 11 a The contour of the number of substep at the 3rd increament, and **b** the strain contour of the foundation (RKF45sec method with STOL = 10^{-9})
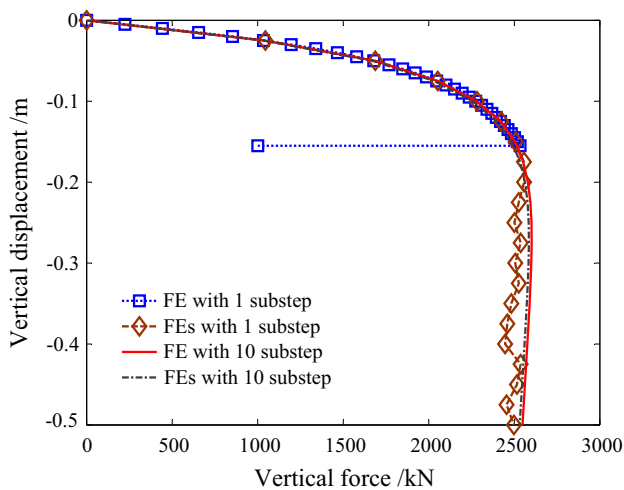

Fig. 12 Relations of vertical force and vertical displacement

Fig. 13 Geometry and boundary conditions of the slope

methods with stress correction treatment. Recently, the shear strength reduction technique is used in SPH to study slope stability with our hypoplastic constitutive model [20].

The geometry and boundary conditions of the considered slope are shown in Fig. 13. The considered slope is discretised by 349 four-node plane strain elements with the bottom of the slope fixed in the horizontal and vertical directions, and the lateral boundaries fixed only in the horizontal direction. The slope is assumed to consist of cohesive soil with the material parameters listed in Table 6, including an initial void ratio of $e_i = 0.88$. The friction angle ϕ and cohesion c are the two shear strength parameters subjected to strength reduction. In the searching process, the actual shear strength is reduced by a factor, that is,

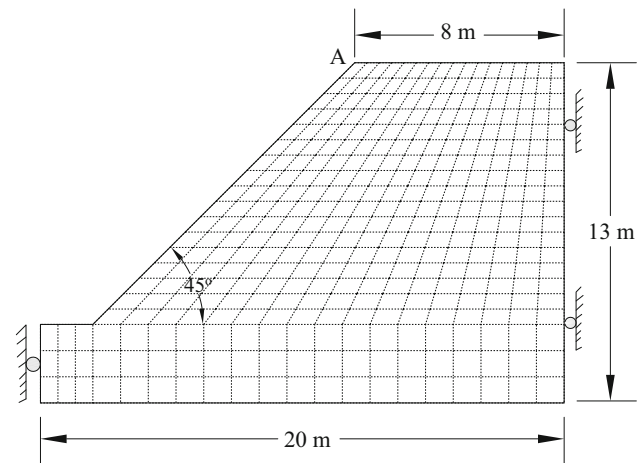$$\phi_f = \frac{\phi}{F_s}, \quad c_f = \frac{c}{F_s}, \tag{42}$$

with the reduced shear strength parameters then used to compute the corresponding hypoplastic parameters C_1, C_2, C_3, C_4 according to the procedure given in [33]. To obtain the initial stress state, a geostatic step is performed by applying gravity loading to the soil. During the geostatic loading, the factor F_s is maintained at a constant value of 0.5 to avoid failure. In the second step, the shear strength parameters are reduced by increasing the factor F_s from 0.5 to 2.0, with the initiation of slope failure defined as occurring at the time when the computation is non-converging. This procedure enabled the adoption of a feasible global increment in the simulations, the numerical results are presented in Table 7.

It can be observed from Table 7 that all the tested explicit methods, with or without stress correction treatment, produced a safety factor of approximately 1.2, corresponding to the shear parameters of about $\phi = 16.7°$ and $c = 10$ kPa. However, the various integration methods

Table 6 Parameters for the homogeneous slope

Parameter	E(Mpa)	v	ϕ	ψ	e_{co}	λ	ξ	α	c (kPa)
Value	100	0.35	20°	0	0.957	0.122	0.061	1.5	12

Table 7 Results of different methods for safety factor of the slope

Method	CPU time (s)	Total Number of increments	Total number of substeps	Max. Number of substeps	Safety factor
FE (100 substeps)	229.2	488	1.396×10^5	48,800	1.2233
FEs (100 substeps)	317.7	482	1.396×10^5	48,200	1.2041
CN method (100 substeps)					
STOL = 10^{-1}	363.5	529	1.396×10^5	42,900	1.284
STOL = 10^{-2}	477.2	559	1.396×10^5	49,300	1.3152
STOL = 10^{-3}	555.8	545	1.396×10^5	54,500	1.2404
STOL = 10^{-4}	316.0	535	1.396×10^5	53,500	1.2938
CNs method (100 substeps)					
STOL = 10^{-1}	Failed at 1st increment				
STOL = 10^{-2}	Failed at 1st increment				
STOL = 10^{-3}	Failed at 1st increment				
STOL = 10^{-4}	Failed at 1st increment				
RKF23sec method					
STOL = 10^{-1}	195.6	547	7,714,296	5526	1.3071
STOL = 10^{-2}	194.5	488	6,908,804	4949	1.2233
STOL = 10^{-3}	221.4	542	7,630,232	6108	1.2992
STOL = 10^{-4}	218.4	531	7,425,739	6050	1.2535
RKF23secs method					
STOL = 10^{-1}	113.0	354	4,943,236	3541	1.0223
STOL = 10^{-2}	135.5	510	7,211,736	5166	1.2563
STOL = 10^{-3}	141.9	510	7,243,457	5966	1.2563
STOL = 10^{-4}	135.0	510	7,308,777	6071	1.2563

exhibited very different levels of performance. Although the forward Euler (FE) and forward Euler with stress correction (FEs) methods both complete around 480 increments of the computation, with every loading increment being divided into 100 equal substeps, the FEs method took approximately 100 s longer than the FE method and also produced a smaller safety factor. In contrast, the implicit CN method with 100 equal substeps took more than 500 increments to obtain the safety factor, while an increase in error tolerance did not result in a corresponding increase in computation time for this method. The CN method with stress correction failed at the first increment, with non-convergence recorded during the local iterations, implying that local iteration is very sensitive to the stress state. Obviously, stress correction may result in a large difference between the current and last substep of the local iterations. Similarly, the adaptive explicit method with or without stress correction required more than 500 increments to obtain the safety factor. Using this method the total number of substeps for all Gauss points is much higher than that for

the FE method and the implicit CN method, although the total CPU time is lower than that required for the other methods without substepping schemes. This further indicates that adaptive explicit methods can effectively save CPU time due to their adaptive nature. Table 7 also reveals that the adaptive explicit method with stress correction, e.g. the RKF23secs method, took less time to complete the calculations than the RKF23sec method without stress correction.

The above results imply that stress correction has a significant influence on numerical computations using hypoplastic models. Figure 14 presents the change in horizontal displacement at the top of the slope (point A in Fig. 13) under different reduction factors and using different integration strategies. It can be observed from this figure that the computations produced using a stress correction scheme recorded horizontal displacements of 0.1 and 0.4 m for the adaptive explicit method and forward Euler method, respectively. In contrast, the same methods without stress correction produced non-convergence with
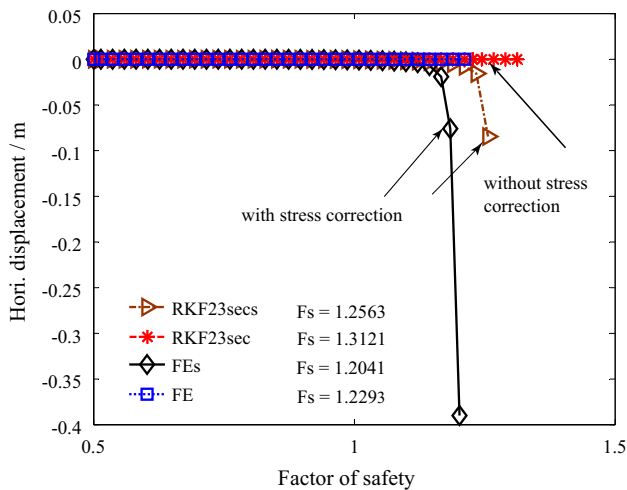
Fig. 14 Change of the horizontal displacement at the top of the slope by strength reduction using different integration strategies

negligible displacement. In addition, there are noticeable difference in the safety factors with and without stress correction (1.26 with correction and 1.31 without correction), which shows the impact of stress correction on the numerical results.

The effect of the stress correction scheme can be further interpreted based on the shear surface of the slope. Figure 15 shows contour plots of the shear surface at the final increment obtained using the RKF23sec method, (I) with and (II) without stress correction. Analysis of this figure reveals that whereas a failure surface, as depicted by the displacement in Fig. 15a, can be observed in the data

produced with a stress correction scheme, no failure surface is generated in the computation undertaken without a stress correction scheme, as shown in Fig. 15b. Similarly, Fig. 15c also shows a shear band (as depicted by the equivalent strain) produced in the computation with stress correction, but no such a failure band is generated in Fig. 15d. Note that the above different results are obtained from computation under different safety factors.

5 Conclusions

This paper presents the numerical implementation of a simple hypoplastic constitutive model using finite element method. Various commonly used integration methods, including both implicit and explicit methods, are enhanced by stress correction, with such influence factors as the load increment size and the specified error tolerance on the performance of the different integration strategies studied. The main conclusions of this work can be summarised as follows:

(1) The hypoplastic model is characterised by a failure surface and a bound surface, which restricts the accessible stresses. For some loading directions, the stress may surpass the failure surface but is still within the bound surface. bound surface largely increases with an increase in the critical friction angle. However, the difference between the failure surface and the bound surface, in term of mobilized
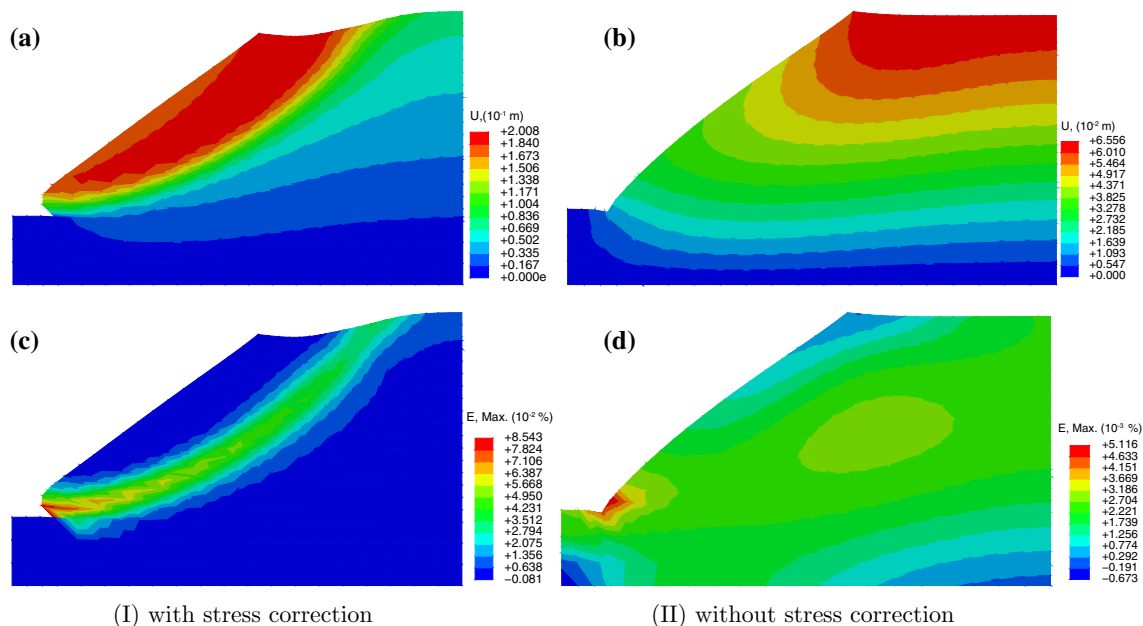

(I) with stress correction

(II) without stress correction

Fig. 15 The displacements (**a**, **b**) and strains (**c**, **d**) upon slope failure by strength reduction using the RKF23sec method, (I) with and (II) without stress correction

friction angle, is so large that a stress correction is necessary.

(2) Compared with the implicit Crank–Nicolson method, the adaptive explicit methods shows better performance in the numerical computations, being less sensitive to the loading direction and increment size since the incremental step size can be changed automatically according to the prescribed accuracy requirement. Moreover, the implicit integration techniques are also sensitive to the specified error tolerance due to the strong nonlinearity of the hypoplastic model. The accuracy can thus be effectively improved by tightening the error tolerance, without increasing the number of substeps nor the CPU time.

(3) Although the tested adaptive explicit methods can produce accurate numerical results, the intrinsic "shortcoming" of the hypoplastic model, that some stresses may lie outside the failure surface, cannot be overcome. In addition, the bound surface is often too far from the failure surface. In order to make sure that the stress does not surpass the failure surface, a stress correction must be employed. Such a stress correction can guarantee that the stress does not go beyond the failure surface. Moreover, the stress correction stabilizes the numerical computation for large increment size.

Appendix

The explicit formula of the failure surface is given in the following form:

$$f(\boldsymbol{\sigma}) = \sqrt{J_2} + \varsigma(C_i, I_e)I_1 = 0,$$

where the implicit variable ς is determined by the dimensionless parameters $C_i(i = 1, 2, 3, 4)$ and the critical state function I_e. The variable ς is expressed as follows:

$$\varsigma = \sqrt{\frac{a - b}{6C_3^2(3C_1^2 - C_4'^2)}},$$

where $C_4' = I_e C_4$ and

$$\begin{aligned}
a = {} & -18C_1^3C_3 + 9C_2^2C_4'^2 + 6C_2C_3C_4'^2 \\
& + C_3^2C_4'^2 + 6C_1(6C_2 + C_3)C_4'^2 \\
& - 6C_1^2(3C_2C_3 + C_3^2 - 6C_4'^2)
\end{aligned}$$

$$\begin{aligned}
b = {} & C_4'(6C_1 + 3C_2 + C_3) \\
& \left\{ -36C_1^3C_3 + 36C_1C_2C_4'^2 \right. \\
& \left. + (3C_2 + C_3)^2C_4'^2 - 36C_1^2(C_2C_3 - C_4'^2) \right\}^{1/2}
\end{aligned}$$

References

1. Abbo AJ (1997) Finite element algorithms for elastoplasticity and consolidation. Ph.D. thesis, University of Newcastle Upon Tyne
2. Artioli E, Auricchio F, Beirao da Veiga L (2007) Generalized midpoint integration algorithms for J2 plasticity with linear hardening. Int J Numer Methods Eng 72(4):422–463
3. Crisfield MA, Remmers JJ, Verhoosel CV (2012) Nonlinear finite element analysis of solids and structures. Wiley, Hoboken
4. Ding YT, Huang WX, Sheng DC, Sloan SW (2015) Numerical study on finite element implementation of hypoplastic models. Comput Geotech 68:78–90
5. Fellin W, Ostermann A (2002) Consistent tangent operators for constitutive rate equations. Int J Numer Anal Meth Geomech 26(12):1213–1233
6. Fellin W, Mittendorfer M, Ostermann A (2009) Adaptive integration of constitutive rate equations. Comput Geotech 36(5):698–708
7. Guo XG, Peng C, Wu W, Wang YQ (2016) A hypoplastic constitutive model for debris materials. Acta Geotech 11(6):1217–1229
8. Heeres OM, de Borst R (2000) Implicit integration of hypoplastic models. In: Kolymbas D (ed) Constitutive modelling of granular materials. Springer, Berlin, pp 457–470
9. Herle I, Kolymbas D (2004) Hypoplasticity for soils with low friction angles. Comput Geotech 31(5):365–373
10. Hill R (1958) A general theory of uniqueness and stability in elastic–plastic solids. J Mech Phys Solids 6(3):236–249
11. Hleibieh J, Wegener D, Herle I (2014) Numerical simulation of a tunnel surrounded by sand under earthquake using a hypoplastic model. Acta Geotech 9(4):631–640
12. Krieg RD, Krieg DB (1997) Accuracies of numerical solution methods for the elastic-perfectly plastic model. J Press Vessel Technol 4(99):510–515
13. Li XS, Wang Y (1998) Linear representation of steady-state line for sand. J Geotech Geoenviron Eng 124(12):1215–1217
14. Li Z, Kotronis P, Escoffier S, Tamagnini S (2016) A hypoplastic macroelement for single vertical piles in sand subject to three-dimensional loading conditions. Acta Geotech 12(2):373–390
15. Lin J, Wu W, Borja RI (2015) Micropolar hypoplasticity for persistent shear band in heterogeneous granular materials. Comput Methods Appl Mech Eng 289(1):24–43
16. Mašín D (2005) A hypoplastic constitutive model for clays. Int J Numer Anal Methods Geomech 29(4):311–336

17. Mašín D, Khalili N (2008) Modelling of the collapsible behaviour of unsaturated soils in hypoplasticity. In: 1st European conference on unsaturated soils. Durham, UK, pp 659–665
18. Niemunis A (2003) Extended hypoplastic models for soils. Habilitation thesis, Ruhr-University
19. Osinov VA, Gudehus G (1996) Plane shear waves and loss of stability in a saturated granular body. Mech Cohes Frict Mater 1:25–44
20. Peng C, Wu W, Yu HS, Wang C (2015) A sph approach for large deformation analysis with hypoplastic constitutive model. Acta Geotech 10(6):703–717
21. Peng C, Guo XG, Wu W, Wang YQ (2016) Unified modelling of granular media with smoothed particle hydrodynamics. Acta Geotech 11(6):1231–1247
22. Potts DM, Gens A (1985) A critical assessment of methods of correcting for drift from the yield surface in elasto-plastic finite element analysis. Int J Numer Anal Methods Geomech 2(9):149–159
23. Rizzi E, Maier G, Willam K (1996) On failure indicators in multi-dissipative materials. Int J Solids Struct 33(20–22):3187–3214
24. Roddeman D (1997) "FEM-implementation of hypoplasticity" European Union project (erbchbgct 940554). Baufakultät, Universität Innsbruck, Inst. für Geotechnik und Tunnelbau
25. Shi XS, Herle I (2017) Numerical simulation of lumpy soils using a hypoplastic model. Acta Geotech 12(2):349–363
26. Sikora Z, Gudehus G (1990) Numerical simulation of penetration in sand based on FEM. Comput Geotech 9(1–2):73–86
27. Sloan SW (1987) Substepping schemes for the numerical integration of elastoplastic stress–strain relations. Int J Numer Methods Eng 24(5):893–911
28. Stutz H, Man D, Wuttke F (2016) Enhancement of a hypoplastic model for granular soilstructure interface behaviour. Acta Geotech 11(6):1249–1261
29. Tamagnini C, Viggiani G, Chambon R (1999) A review of two different approaches to hypoplasticity. In: Kolymbas D (ed) Constitutive modelling of granular materials. Springer, Berlin, pp 107–145
30. Tamagnini C, Viggiani G, Chambon R, Desrues J (2000) Evaluation of different strategies for the integration of hypoplastic constitutive equations. Application to the CLoE model. Mech Cohes Frict Mater 5(4):263–289
31. Tejchman J, Wu W (1996) Numerical simulation of shear band formation with a hypoplastic constitutive model. Comput Geotech 18(1):71–84
32. Von Wolffersdorff PA (1996) A hypoplastic relation for granular materials with a predefined limit state surface. Mech Cohes Frict Mater 1(3):251–271
33. Wang XT, Wu W (2011) An updated hypoplastic constitutive model, its implementation and application. In: Wan R, Alsaleh M, Labuz J (eds) Bifurcations, instabilities and degradations in geomaterials. Springer, Berlin, pp 133–143
34. Wissmann JW, Hauck C (1983) Efficient elastic–plastic finite element analysis with higher order stress-point algorithms. Comput Struct 17(1):89–95
35. Wu W (1990) A unified numerical integration formula for the perfectly plastic von mises model. Int J Numer Methods Eng 30(3):491–504
36. Wu W (1999) On a simple critical state model for sand. In: Proceedings of the 7th international symposium numerical models in geomechanics, pp 47–52
37. Wu W, Bauer E (1994) A simple hypoplastic constitutive model for sand. Int J Numer Anal Methods Geomech 18(12):833–862
38. Wu W, Kolymbas D (1990) Numerical testing of the stability criterion for hypoplastic constitutive equations. Mech Mater 9(3):245–253
39. Wu W, Niemunis A (1996) Failure criterion, flow rule and dissipation function derived from hypoplasticity. Mech Cohes Frict Mater 1(2):145–163
40. Wu W, Niemunis A (1997) Beyond failure in granular materials. Int J Numer Anal Methods Geomech 21(3):153–174
41. Wu W, Bauer E, Kolymbas D (1996) Hypoplastic constitutive model with critical state for granular materials. Mech Mater 23(1):45–69
42. Wu W, Lin J, Wang X (2017) A basic hypoplastic constitutive model for sand. Acta Geotech 12(6):1373–1382