Computers and Geotechnics 37 (2010) 59–67

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Explicit stress integration with error control for the Barcelona Basic Model Part I: Algorithms formulations

Wojciech T. Sołowski [a,*,1], Domenico Gallipoli [b]

[a] The University of Newcastle, Callaghan NSW 2308, Australia
[b] University of Glasgow, Glasgow G12 8LT, UK

ARTICLE INFO

Article history:
Received 18 January 2009
Received in revised form 9 June 2009
Accepted 13 July 2009
Available online 8 August 2009

Keywords:
Explicit stress integration
Richardson extrapolation
Runge–Kutta methods
Unsaturated soils
Barcelona Basic Model
Elasto-plasticity

ABSTRACT

The numerical integration of the stress–strain relationship is an important part of many finite element code used in geotechnical engineering. The integration of elasto-plastic models for unsaturated soils poses additional challenges associated to the presence of suction as an extra constitutive variable with respect to traditional saturated soil models. In this contribution, a range of explicit stress integration schemes are derived with specific reference to the Barcelona Basic Model (BBM), which is one of the best known elasto-plastic constitutive models for unsaturated soils. These schemes, however, do not address possible non-convexity of the loading collapse (LC) curve and neglect yielding on the suction increase (SI) line. The paper describes eight Runge–Kutta methods of various orders with adaptive substepping as well as a novel integration scheme based on Richardson extrapolation. The algorithms presented also incorporate two alternative error control methods to ensure accuracy of the numerical integration. Extensive validation and comparison of different schemes are presented in a companion paper. Although the algorithms presented were coded for the Barcelona Basic Model, they can be easily adapted to other unsaturated elasto-plastic models formulated in terms of two independent stress variables such as net stress and suction.

1. Introduction

Finite element codes used for nonlinear analyses in geotechnical engineering employ iterative schemes, which require solving the global system of equations several times for each load/time increment. In each such iteration, the stress state at every Gauss point is calculated from the corresponding strain increment using a stress integration procedure. The numerical integration of the stress–strain relationship is therefore one of the most frequent tasks performed by a finite element routine and, for complex elasto-plastic constitutive models formulated in a differential form, the choice of the integration algorithm directly influences the stability, efficiency and accuracy of computations [14].

The algorithms used for the numerical integration of elasto-plastic constitutive models may be divided into two general groups: (a) the implicit algorithms where the tangent elasto-plastic stiffness matrix is computed corresponding to the final stress state of the integration increment (hence, these algorithms require an iterative calculation of the final stress state) and (b) the explicit algorithms where the tangent elasto-plastic stiffness matrix is computed corresponding to the initial stress state of the integration increment. Unlike implicit algorithms, the explicit algorithms do not require an iterative calculation but they tend to use adaptive substepping to ensure accuracy and to control error.

A general description of implicit integration algorithms is given e.g. in Simo and Hughes [18] where references to related works may also be found. In the relatively new area of unsaturated soil mechanics, implicit integration algorithms have been presented e.g. by Vaunat et al. [26], Zhang et al. [28] and Borja [3].

As mentioned earlier, explicit integration algorithms are often used in combination with an adaptive substepping procedure to control accuracy of results. The substepping procedure divides the initial strain increment into a number of substeps based on a particular criterion to ensure that the integration error is maintained below a given tolerance while keeping the integration as effective as possible. These concepts introduced to geomechanics by Sloan [19] are typical of modern explicit stress integration algorithms. Recent developments in explicit integration have been presented by Sloan et al. [20], Sheng et al. [14–17] and Zhao et al. [29], whereas a general summary of integration algorithms for soils with a comparison between explicit and implicit approaches may be found in Potts and Zdravkovic [11].

This paper, together with a companion one [24], compares the robustness, accuracy and efficiency of different Runge–Kutta algorithms for the integration of an elasto-plastic constitutive model

* Corresponding author. Tel.: +61 2 49217074; fax: +61 2 4921 6991.
E-mail addresses: wojteks@gmx.net, wojciech.solowski@newcastle.edu.au (W.T. Sołowski).
[1] Formerly Durham University, UK.

0266-352X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compgeo.2009.07.004

for unsaturated soils, expanding on the original idea of explicit integration with adaptive substepping proposed by Sloan [19]. A novel explicit stress integration algorithm based on the idea of Richardson extrapolation is also described and compared with the Runge–Kutta schemes. All algorithms are presented with specific reference to the well-known Barcelona Basic Model (BBM) proposed by Alonso et al. [1], though the version of BBM considered in this work neglects the possibility of yielding on the suction increase (SI) line. This simplifies treatment with little loss of generality as, for most applications of BBM, soil paths never intersect the SI locus. In addition, the loading collapse (LC) yield locus may be non-convex for some specific set of parameter values [27], however this issue is not specifically addressed in the algorithms presented here.

Although the numerical examples in both this paper and the companion one [24] refer specifically to BBM, some of the conclusions may be generally applicable to other types of elasto-plastic unsaturated soil models.

2. Barcelona Basic Model

The BBM has been the first elasto-plastic model to offer a comprehensive interpretation of the irreversible deformation of unsaturated soils as a consequence of both loading and suction changes. Partial saturation is taken into account in BBM by including an additional scalar variable, i.e. the suction $s = u^{(a)} - u^{(w)}$, defined as the difference between the pore air pressure $u^{(a)}$ and pore water pressure $u^{(w)}$. For the particular case of zero suction, the BBM reduces to the Modified Cam Clay for saturated soils. The BBM was originally formulated by Alonso et al. [1] for the $p - q$ triaxial net stress space, where p is the mean net stress, defined as the excess of the mean total stress over pore air pressure, and q is the triaxial deviator stress. The BBM assumes that the size of the yield locus in the $p - q$ net stress plane at constant suction is dependent on the current values of suction s and hardening parameter p_0^*.

The BBM is calibrated using nine material constants, namely the elastic shear modulus G, the elastic swelling modulus κ for changes in mean net stress, the elastic swelling modulus κ_s for changes in suction, the parameter k describing the increase of cohesion with suction, the slope of the critical state line M in the $q - p$ plane at constant suction, the slope of the normal consolidation line $\lambda(0)$ at zero suction, the parameter r defining the range of slopes for the constant suction normal consolidation lines, the parameter β controlling the change of the slope of the constant suction normal consolidation lines with suction and the reference mean net stress p^c coinciding with the value of the hardening parameter p_0^* for which the isotropic yield stress is independent of suction.

The value of the specific volume at a reference net stress state must also be specified as an input parameter. This information is usually provided in the form of the parameter $N(0)$ representing the specific volume on the normal consolidation line at zero suction and at the reference mean net stress p^c.

For the application to generic boundary value problems, BBM must be recast in the general net stress space where the net stress vector $\boldsymbol{\sigma}$ is defined as:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{tot} - \mathbf{m}\, u^{(a)} \tag{1}$$

where $\boldsymbol{\sigma}^{tot}$ is the total stress vector and $\mathbf{m}^T = \{1,1,1,0,0,0\}$. The mean net stress p and triaxial shear stress q are, respectively:

$$p = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33})_1 \tag{2}$$

$$q = \sqrt{\frac{1}{2}((\sigma_{11} - \sigma_{22})^2 + (\sigma_{11} - \sigma_{33})^2 + (\sigma_{22} - \sigma_{33})^2 + 6(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2))} \tag{3}$$

No dependency of the yield locus on the Lode angle is assumed (i.e. each constant suction yield locus has a circular cross-section in the deviatoric plane) and the original BBM formulation in the triaxial net stress space is therefore extended to the general stress domain by simply defining p and q in terms of Eqs. (2) and (3), respectively.

3. Normalised yield locus

The definition of a normalised form of the yield equation for BBM is necessary to set a stress integration tolerance that is independent of the current size of the yield locus. In a constant suction $p - q$ net stress plane, the BBM yield locus is given by an ellipse, which size depends on the current values of suction s and hardening parameter p_0^* according to the equation:

$$F = q^2 - M^2(p + ks)(p_0 - p) = 0 \tag{4}$$

where p_0 is the isotropic yield stress depending on the current values of s and p_0^*. A normalised form of the yield locus is obtained by dividing (4) by the quantity $(p_0 + ks)^2$

$$F_{norm} = q'^2 - M^2 p'(1 - p') = 0 \tag{5}$$

where the dependency on the two dimensionless variables $q' = \frac{q}{p_0 + ks}$ and $p' = \frac{p + ks}{p_0 + ks}$ is introduced.

Regardless of the current values of suction, s and hardening parameter, p_0^*, the normalised Eq. (5) maps all the BBM yield loci given by Eq. (4) into a single ellipse in the p'–q' plane. The major and minor axes of this ellipse are equal to unity and half the value of M, respectively. The value of the normalised yield function F_{norm} depends on the distance of the current net stress state from the yield locus of Eq. (4) relative to the size of such yield locus (which in turn depends on the current values of suction s and hardening parameter p_0^*).

In all stress integration algorithms presented in this paper, the net stress state is assumed to lie on the yield locus when the corresponding absolute value of the normalised yield function F_{norm} of Eq. (5) is less then a fixed tolerance. Note that, if Eq. (4) was to be used instead of Eq. (5) in conjunction with an absolute tolerance, the yield criterion will become less stringent as the size of the yield locus decreases. In particular, for a very small size of the yield locus, the net stress state might be far from the yield surface in relative terms but might still lie within the set absolute tolerance.

4. Explicit stress integration of unsaturated soil models with substepping

Sheng et al. [15] proposed an explicit substepping stress integration algorithm for an elasto-plastic unsaturated soil model formulated in terms of a constitutive stress variable defined as:

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma}^{tot} - \mathbf{m}\varphi(S_r)u^{(w)} \tag{6}$$

where $\boldsymbol{\sigma}'$ is the constitutive stress vector, $\varphi(S_r)$ is a function of the degree of saturation S_r and $u^{(w)}$ is the pore water pressure.

Expanding on the work of Sheng et al. [15], this paper compares different Runge–Kutta schemes of various orders used for the explicit integration of BBM in combination with an adaptive substepping procedure for error control. In these algorithms, the change of suction is treated as an additional component of the strain increment vector despite, in the context of constitutive modelling, suction is considered as a stress rather than a strain. A mixed "strain–suction" formulation is here adopted because suction, similarly to strains, is derived from the primary nodal variables (i.e. displacements, pore air pressure and pore water pressure) of finite element models for unsaturated soils. Hence, the stress integration algorithm should return the increment of net stress $\Delta\boldsymbol{\sigma}$ corresponding

to an "enhanced" vector increment $\Delta\boldsymbol{\varepsilon}^{enh} = \{\delta\boldsymbol{\varepsilon}, \Delta s\}$, which includes the conventional strain increment vector $\Delta\boldsymbol{\varepsilon}$ as well as the suction increment Δs.

The substepping procedure automatically divides the "enhanced" increment of strains and suction into a number of substeps $\delta\boldsymbol{\varepsilon}^{enh} = \{\delta\boldsymbol{\varepsilon}, \delta s\}$ small enough to ensure that the desired integration accuracy is enforced. The integration error in the current substep is assessed and, depending on the error control criterion, the substep is either accepted or rejected. After accepting or rejecting the current substep, the size of the next substep is calculated based on the estimated error and the set tolerance.

Two different criteria for controlling the integration error are here evaluated: (i) the EPS (Error Per Step) control, as used e.g. by Sloan [19], and (ii) the EPUS (Error Per Unit Step), as defined by Shampine [13]. To the authors' knowledge, the EPUS error control has not been used so far in adaptive stress integration of soil models.

The stress integration algorithm consists of two parts, i.e. an elastic part and an elasto-plastic part, which are alternatively used depending on whether the given increment of strains and suction corresponds to elastic or elasto-plastic loading, respectively. In the particular case of an increment of strains and suction starting as elastic and arriving to a plastic condition, the algorithm must separate the initial portion of the increment that takes place inside the yield locus from the remaining part that involves plastic behaviour. For the sake of brevity, this paper does not cover elastic integration of stresses nor the computational procedures used for dividing an increment in the corresponding elastic and elasto-plastic parts (for details about such aspects, the reader should refer to e.g. Potts and Zdravkovic [11] and Sołowski [21]). This paper focuses instead on the stress integration in the plastic domain and, hence, the derivation presented in the next section assumes that the soil behaves elasto-plastically right from the onset of the integration.

With specific reference to BBM, the differential form of the consistency condition is written as:

$$dF = \left(\frac{\partial F}{\partial \boldsymbol{\sigma}}\right)^T d\boldsymbol{\sigma} + \left(\frac{\partial F}{\partial s}\right)^T ds + \left(\frac{\partial F}{\partial p_0^*}\right)^T dp_0^* = 0 \tag{7}$$

The stress increment $d\boldsymbol{\sigma}$ in (7) is computed from the integration of the elastic strain associated to changes of net stress, $d\boldsymbol{\varepsilon}^{el,\sigma}$ (given by the difference between the strain $d\boldsymbol{\varepsilon}$ and the sum of the plastic strain $d\boldsymbol{\varepsilon}^{pl}$ plus the elastic strain associated to changes of suction $d\boldsymbol{\varepsilon}^{el,s}$):

$$d\boldsymbol{\sigma} = \mathbf{D}^{el} d\boldsymbol{\varepsilon}^{el,\sigma} = \mathbf{D}^{el}[d\boldsymbol{\varepsilon} - (d\boldsymbol{\varepsilon}^{pl} + d\boldsymbol{\varepsilon}^{el,s})] \tag{8}$$

where $\mathbf{D}^{el}$ is the elastic matrix.

The elastic strain associated to changes of suction is defined as:

$$d\boldsymbol{\varepsilon}^{el,s} = \frac{1}{3}\mathbf{m}\frac{\kappa_s ds}{(p^{at}+s)v} \tag{9}$$

where κ_s is the elastic swelling modulus for changes in suction, p^{at} is the atmospheric pressure and v is the current specific volume.

The plastic strain increment is computed from the plastic potential function as:

$$d\boldsymbol{\varepsilon}^{pl} = \lambda\frac{\partial Q}{\partial \boldsymbol{\sigma}} \tag{10}$$

where λ is the scalar plastic multiplier and $Q = Q(\boldsymbol{\sigma}, s, p_0^*)$ is the plastic potential function. Similarly, the increment of the hardening parameter dp_0^* is computed from the increment of plastic strain as:

$$dp_0^* = \frac{\partial p_0^*}{\partial \varepsilon_v^{pl}} d\varepsilon_v^{pl} = \frac{\partial p_0^*}{\partial \varepsilon_v^{pl}}\mathbf{m}^T d\boldsymbol{\varepsilon}^{pl} \tag{11}$$

where $d\varepsilon_v^{pl}$ is the increment of plastic volumetric strain. Introducing (8)–(11) into (7) leads to the following expression for the scalar plastic multiplier λ:

$$\lambda = \frac{\mathbf{a}^T\mathbf{D}^{el}d\boldsymbol{\varepsilon} + [c - \mathbf{a}^T\mathbf{D}^{el}\mathbf{b}]ds}{\mathbf{a}^T\mathbf{D}^{el}\mathbf{g} - d} \tag{12}$$

where

$$\mathbf{a} = \frac{\partial F}{\partial \boldsymbol{\sigma}}, \quad \mathbf{b} = \frac{1}{3}\mathbf{m}\frac{\kappa_s}{(p^{at}+s)v}, \quad c = \frac{\partial F}{\partial s}, \quad d = \frac{\partial F}{\partial p_0^*}\frac{\partial p_0^*}{\partial \varepsilon_v^{pl}}\mathbf{m}^T\frac{\partial Q}{\partial \boldsymbol{\sigma}} \quad \text{and} \quad \mathbf{g} = \frac{\partial Q}{\partial \boldsymbol{\sigma}}$$

The plastic strain increment is then calculated from (10), the hardening parameter increment from (11) and the net stress increment from (8). In the calculation of the net stress increment, the terms associated with strain and suction can be separated, which leads to:

$$d\boldsymbol{\sigma} = \mathbf{D}^{ep,\varepsilon}d\boldsymbol{\varepsilon} + \mathbf{D}^{ep,s}ds = [\mathbf{D}^{ep,\varepsilon}, \mathbf{D}^{ep,s}]d\boldsymbol{\varepsilon}^{enh} = \mathbf{D}^{ep}d\boldsymbol{\varepsilon}^{enh} \tag{13}$$

where the $\mathbf{D}^{ep}$ is the 6×7 tangent elasto-plastic matrix given by the combination of the 6×6 matrix $\mathbf{D}^{ep,\varepsilon}$ and the 6×1 matrix $\mathbf{D}^{ep,s}$ as:

$$\mathbf{D}^{ep} = \left[\underbrace{\begin{matrix}\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\end{matrix}}_{\mathbf{D}^{ep,\varepsilon}} \middle| \underbrace{\begin{matrix}\cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots\end{matrix}}_{\mathbf{D}^{ep,s}}\right] \tag{14}$$

where $\mathbf{D}^{ep,\varepsilon}$ and $\mathbf{D}^{ep,s}$ are, respectively, defined as:

$$\mathbf{D}^{ep,\varepsilon} = \mathbf{D}^{el} - \frac{\mathbf{D}^{el}\mathbf{g}\mathbf{a}^T\mathbf{D}^{el}}{\mathbf{a}^T\mathbf{D}^{el}\mathbf{g} - d} \tag{15}$$

and

$$\mathbf{D}^{ep,s} = \mathbf{D}^{el}\mathbf{g}\frac{(c - \mathbf{a}^T\mathbf{D}^{el}\mathbf{b})}{\mathbf{a}^T\mathbf{D}^{el}\mathbf{g} - d} - \frac{1}{3}\mathbf{D}^{el}\mathbf{m}\frac{\kappa_s}{(p^{at}+s)v} \tag{16}$$

Similarly, the hardening parameter increment is calculated by introducing Eqs. (10) and (12) into Eq. (11), with the terms associated with strains and suction separated as:

$$dp_0^* = \mathbf{H}^{\varepsilon}d\boldsymbol{\varepsilon} + \mathbf{H}^{s}ds = [\mathbf{H}^{\varepsilon}, \mathbf{H}^{s}]d\boldsymbol{\varepsilon}^{enh} = \mathbf{H}d\boldsymbol{\varepsilon}^{enh} \tag{17}$$

where $\mathbf{H}$ is the 1×7 hardening matrix given by the combination of the 1×6 matrix $\mathbf{H}^{\varepsilon}$ and the 1×1 matrix $\mathbf{H}^{s}$ as:

$$\mathbf{H} = \left[\underbrace{\begin{matrix}\cdots & \cdots & \cdots & \cdots & \cdots & \cdots\end{matrix}}_{\mathbf{H}^{\varepsilon}} \middle| \underbrace{\begin{matrix}\cdots\end{matrix}}_{\mathbf{H}^{s}}\right] \tag{18}$$

where $\mathbf{H}^{\varepsilon}$ and $\mathbf{H}^{s}$ are, respectively, defined as:

$$\mathbf{H}^{\varepsilon} = \frac{\partial p_0^*}{\partial \varepsilon_v^{pl}}\frac{\mathbf{m}^T\mathbf{g}\mathbf{a}^T\mathbf{D}^{el}}{\mathbf{a}^T\mathbf{D}^{el}\mathbf{g} - d} \tag{19}$$

and

$$\mathbf{H}^{s} = \frac{\partial p_0^*}{\partial \varepsilon_v^{pl}}\frac{\mathbf{m}^T\mathbf{g}[c - \mathbf{a}^T\mathbf{D}^{el}\mathbf{b}]}{\mathbf{a}^T\mathbf{D}^{el}\mathbf{g} - d} \tag{20}$$

In the adaptive substepping procedure the infinitesimal increments of Eq. (13) (denoted by "d") are replaced by the corresponding finite substeps (denoted by "δ"). The sum of all substeps $\delta\boldsymbol{\sigma}$ of net stress and the sum of all substeps $\delta\boldsymbol{\varepsilon}^{enh}$ of strains and suction add up, over a given increment, to the corresponding values $\Delta\boldsymbol{\sigma}$ and $\Delta\boldsymbol{\varepsilon}^{enh}$, which are the output and input of the integration algorithm, respectively.

The magnitude of the generic ith substep of a given increment of strains and suction is therefore defined as:

$$\delta\boldsymbol{\varepsilon}_i^{enh} = \delta T_i \cdot \Delta\boldsymbol{\varepsilon}^{enh} \tag{21}$$

where δT_i is the substep fraction smaller than unity. Assuming that the initial strain increment has been partitioned in n substeps, we have that:

$$\sum \delta T_i = 1 \quad \text{for } i = 1, \ldots, n \tag{22}$$

5. Higher order Runge–Kutta schemes

Different Runge–Kutta schemes have been used for the explicit integration of net stresses within each substep. Runge–Kutta schemes with embedded pairs provide two solutions (i.e. a higher order solution as well as an embedded lower order solution) so that an estimate of the integration error can be easily obtained as the difference between the pair. These Runge–Kutta schemes have been here used in the extrapolation mode, which means that the higher order solution is used to advance the integration over the increment. The following explicit schemes have been investigated in this work:

- The Modified Euler 2(1) scheme (Sloan [19]) – a second order scheme with embedded first order solution.
- Two different 3(2) third order schemes with embedded second order solution. The first scheme uses coefficients proposed by Nystrom (Gear [6] and Lee and Schiesser [9]) whereas the second scheme uses the coefficients proposed by Bogacki–Shampine (Bogacki and Shampine [2]).
- A 4(3) fourth order scheme with embedded third order solution (Lee and Schiesser [9]).
- Four different 5(4) fifth order schemes with embedded fourth order solution. The coefficients for each of these four schemes were proposed by England (Sloan [19] and Lee and Schiesser [9]), Cash and Karp (Press et al. [12] and Lee and Schiesser [9]), Dormand and Prince (Dormand [5]) and Bogacki and Shampine (Bogacki and Shampine [2]), respectively.

The integration of a given substep of strains and suction $\delta\boldsymbol{\varepsilon}^{enh}$ is performed by any of the above Runge–Kutta schemes following the same general procedure [22]. The reader is directed to specialized references for a detailed description of the Runge–Kutta method (e.g. [4,5,13]). Here it suffices to say that, for a given Runge–Kutta scheme, a number (NoS) of stages is necessary to integrate the net stress state over the given substep $\delta\boldsymbol{\varepsilon}^{enh}$. In the generic jth stage (with $j = 1, \ldots, \text{NoS}$), the stage-substeps of net stress $\delta\boldsymbol{\sigma}^{(j)}$ and hardening parameter $\delta p_0^{*(j)}$ are calculated from (13) and (17), respectively, as:

$$\begin{aligned}
\delta\boldsymbol{\sigma}^{(j)} &= \mathbf{D}^{ep(j)}\left(\boldsymbol{\varepsilon}^{enh} + c^{(j)}\delta\boldsymbol{\varepsilon}^{enh},\quad \boldsymbol{\sigma} + \sum_{k=1}^{j-1} a^{(jk)}\delta\boldsymbol{\sigma}^{(k)},\right.\\
&\qquad \left. p_0^* + \sum_{k=1}^{j-1} a^{(jk)}\delta p_0^{*(k)}\right)\delta\boldsymbol{\varepsilon}^{enh}\\
\delta p_o^{*(j)} &= \mathbf{H}^{(j)}\left(\boldsymbol{\varepsilon}^{enh} + c^{(j)}\delta\boldsymbol{\varepsilon}^{enh},\quad \boldsymbol{\sigma} + \sum_{k=1}^{j-1} a^{(jk)}\delta\boldsymbol{\sigma}^{(k)},\right.\\
&\qquad \left. p_0^* + \sum_{k=1}^{j-1} a^{(jk)}\delta p_0^{*(k)}\right)\delta\boldsymbol{\varepsilon}^{enh}
\end{aligned} \tag{23}$$

In Eq. (23) the tangent elasto-plastic matrix $\mathbf{D}^{ep(j)}$ and the hardening matrix $\mathbf{H}^{(j)}$ for the generic jth stage, are a function of the previous stage-substeps of net stress $\delta\boldsymbol{\sigma}^{(k)}$ and hardening parameter $\delta p_o^{*(k)}$ as well as of the substep of strains and suction $\delta\boldsymbol{\varepsilon}^{enh}$. The coefficients $a^{(jk)}$ and $c^{(j)}$ in Eq. (23) are relevant to the specific Runge–Kutta scheme considered. Finally, after all (NoS) stage-substeps have been calculated from Eq. (23), the substeps of net stress $\delta\boldsymbol{\sigma}$ and hardening parameter $\delta p_o^{*(k)}$ are obtained from the weighted summation of such stage-substeps as:

$$\begin{aligned}
\delta\boldsymbol{\sigma} &= \sum_{j=1}^{\text{NoS}} b^{(j)}\delta\boldsymbol{\sigma}^{(j)}\\
\delta p_o^* &= \sum_{j=1}^{\text{NoS}} b^{(j)}\delta p_o^{*(j)}
\end{aligned} \tag{24}$$

where $b^{(j)}$ are coefficients relevant to the specific Runge–Kutta scheme.

Similarly, the weighted summation for the embedded lower order formula gives:

$$\begin{aligned}
\delta\boldsymbol{\sigma}^{(low)} &= \sum_{j=1}^{\text{NoS}} d^{(j)}\delta\boldsymbol{\sigma}^{(j)}\\
\delta p_o^{*(low)} &= \sum_{j=1}^{\text{NoS}} d^{(j)}\delta p_0^{*(j)}
\end{aligned} \tag{25}$$

where $d^{(j)}$ are coefficients relevant to the specific Runge–Kutta scheme.

The estimates of the errors in the calculation of the substeps of net stress and hardening parameter are then given as:

$$\begin{aligned}
\mathbf{E}(\delta\boldsymbol{\sigma}) &= \delta\boldsymbol{\sigma} - \delta\boldsymbol{\sigma}^{(low)} = \sum_{j=1}^{\text{NoS}} e^{(j)}\delta\boldsymbol{\sigma}^{(j)} = \sum_{j=1}^{\text{NoS}} (b^{(j)} - d^{(j)})\delta\boldsymbol{\sigma}^{(j)}\\
E(\delta p_0^*) &= \delta p_o^* - \delta p_o^{*(low)} = \sum_{j=1}^{\text{NoS}} e^{(j)}\delta p_o^{*(j)} = \sum_{j=1}^{\text{NoS}} (b^{(j)} - d^{(j)})\delta p_o^{*(j)}
\end{aligned} \tag{26}$$

The coefficients $a^{(jk)}$, $b^{(j)}$, $c^{(j)}$, $d^{(j)}$ and $e^{(j)}$ for all Runge–Kutta schemes considered in this work are given in Appendix A.

Two additional error estimates are calculated relative to the substeps of mean net stress and deviator stress as:

$$\begin{aligned}
E(\delta p) &= (E(\delta\sigma_{11}) + E(\delta\sigma_{22}) + E(\delta\sigma_{33}))/3\\
E(\delta q) &= \max\{|q(\boldsymbol{\sigma} + \delta\boldsymbol{\sigma} + \mathbf{E}(\delta\boldsymbol{\sigma})) - q(\boldsymbol{\sigma} + \delta\boldsymbol{\sigma})|,\\
&\qquad |q(\boldsymbol{\sigma} + \delta\boldsymbol{\sigma} - \mathbf{E}(\delta\boldsymbol{\sigma})) - q(\boldsymbol{\sigma} + \delta\boldsymbol{\sigma})|\}
\end{aligned} \tag{27}$$

The error estimates of Eqs. (26) and (27) are used to decide whether the substep should be accepted or rejected and to calculate the next substep size by using either the EPS or the EPUS control method as described below.

6. The EPS control method

According to the original substepping procedure proposed by Sloan [19], a relative measure of the error in net stress $\mathbf{R}(\delta\boldsymbol{\sigma})$ is calculated at end of the generic substep as:

$$\mathbf{R}(\delta\boldsymbol{\sigma}) = \frac{\mathbf{E}(\delta\boldsymbol{\sigma})}{|\boldsymbol{\sigma} + \delta\boldsymbol{\sigma}|} \tag{28}$$

In this work, additional relative errors of deviator stress, $R(\delta q)$, mean net stress, $R(\delta p)$ and hardening parameter, $R(\delta p_o^{*(k)})$ are calculated at end of the generic substep as:

$$R(\delta p) = \frac{E(\delta p)}{p + \delta p},\quad R(\delta q) = \frac{E(\delta q)}{q + \delta q},\quad R(\delta p_o^*) = \frac{E(\delta p_0^*)}{p_0^* + \delta p_0^*} \tag{29}$$

The maximum value, R is then identified from the six components of the relative error vector of Eq. (28) and the three scalar errors of Eq. (29):

$$R = max(|R(\delta\sigma_{11})|,\ |R(\delta\sigma_{22})|,\ |R(\delta\sigma_{33})|,\ |R(\delta\sigma_{12})|,\ |R(\delta\sigma_{13})|,\ |R(\delta\sigma_{23})|,\ |R(\delta p)|,\ |R(\delta q)|,\ |R(\delta p_0^*)|) \tag{30}$$

The criterion for error control compares the value of R from Eq. (30) with a user-defined integration tolerance SSTOL. If $R \leqslant$ SSTOL, the substep is accepted and the stress state is updated accordingly whereas if $R >$ SSTOL, the substep is rejected and the net stress state remains unchanged to the value calculated at the end of the previous substep.

Regardless of whether the current substep is accepted or rejected, a new substep size is calculated by defining an updated substep fraction δT_{new} as:

$$\delta T_{new} = \chi \delta T_i \tag{31}$$

where δT_i is the current substep fraction (see Eq. (21)) and χ is a coefficient, which is usually smaller than one if the current substep has been rejected and larger than one if the current substep has been accepted. The coefficient χ in Eq. (31) is calculated as:

$$\chi = \xi \sqrt[m]{\frac{\text{SSTOL}}{R}} \tag{32}$$

where m is the order of the Runge–Kutta method and ξ is a user-defined factor, smaller than one, introduced because of the simplifying assumptions made during the calculation of χ. In this way, the substep size is continuously adapted (i.e. reduced or enlarged depending on whether the current substep has been rejected or accepted) to enforce that the magnitude of the maximum relative error R is maintained close to SSTOL. This approach is often referred to as the "Error Per Step" (EPS) control method (see Shampine [13]).

7. The EPUS control method

In this work, the following definitions of the relative integration error, which are alternative to those given by Eqs. (28) and (29), have also been used:

$$R(\delta\sigma_{ij}) = \frac{E(\delta\sigma_{ij})}{\delta\sigma_{ij}} \quad i,j = 1,2,3 \tag{33}$$

$$R(\delta p) = \frac{E(\delta p)}{\delta p}, \quad R(\delta q) = \frac{E(\delta q)}{\delta q}, \quad R(\delta p_0^*) = \frac{E(\delta p_0^*)}{\delta p_0^*} \tag{34}$$

The maximum relative error R and the updated substep fraction δT_{new} are calculated in the same way as described in the previous section, with the exception of Eq. (32) that is now replaced by:

$$\chi = \xi \sqrt[m-1]{\frac{\text{SSTOL}}{R}} \tag{35}$$

The reduction by one unit of the root order in Eq. (35) in comparison to Eq. (32) has a theoretical justification (see e.g. Shampine [13]). The numerical tests performed in this work and presented in the companion paper [24] have, however, rarely shown a significant difference in efficiency, accuracy or stability if Eq. (32) is used instead of the mathematically correct Eq. (35).

The present approach is often referred to as the "Error Per Unit Step" (EPUS) control method (see also Shampine [13]) and can be used as an alternative to the EPS control method described in the previous section. Note that the EPUS control method has not been previously used for the explicit integration of constitutive models for soils.

Comparison of Eqs. (28) and (29) with Eqs. (33) and (34) indicates that the EPUS control method calculates the error *relative to the net stress change* rather than *relative to the absolute net stress magnitude*, as in the EPS control method. This difference becomes relevant when the integration error is negligible with respect to the absolute magnitude of net stress but significant with respect to the calculated net stress increment. In this circumstance, the error control in terms of Eqs. (28) and (29) would be satisfied whereas the error control in terms of Eqs. (33) and (34) would fail. It therefore seems intuitive that the error definitions of Eqs. (33) and (34) should be favoured over the error definitions of Eqs. (28) and (29) if the error must be kept below a certain percentage of the net stress increment. However, this intuitive conclusion has not been entirely confirmed by the results obtained in this work. As shown by the numerical tests presented in the companion paper [24], the use of Eqs. (28) and (29) will result most of the times (although not always) in integration errors – calculated as relative to the increment of net stress – that are lower than the tolerance SSTOL. Such behaviour is characteristic for strictly dissipative systems as shown in Deuflhard and Bornemann [4].

8. Drift correction scheme

At the end of each plastic substep the stress state should lie on the normalised yield locus given by Eq. (5) within a fixed tolerance. If it does not, a drift correction algorithm is employed, similar to the one advised by Potts and Gens [10] for saturated elasto-plastic models. The algorithm calculates a correction of the net stress state $\delta\boldsymbol{\sigma}^{(corr)}$ while the increment of strains and suction are kept unchanged (i.e. no variation of either strains or suction occur during correction of the net stress state). This implies that the correction of the elastic strains due to change in net stress $\delta\boldsymbol{\varepsilon}^{el,\sigma(corr)}$ must be equal in absolute value and opposite in sign to the correction of the plastic strains $\delta\boldsymbol{\varepsilon}^{pl(corr)}$ and, hence, the correction of the net stress state $\delta\boldsymbol{\sigma}^{(corr)}$ is calculated as:

$$\delta\boldsymbol{\sigma}^{(corr)} = \mathbf{D}^{el}\delta\boldsymbol{\varepsilon}^{el,\sigma(corr)} = -\mathbf{D}^{el}\delta\boldsymbol{\varepsilon}^{pl(corr)} = -\lambda^{(corr)}\mathbf{D}^{el}\frac{\partial Q}{\partial\boldsymbol{\sigma}} \tag{36}$$

where the scalar multiplier $\lambda^{(corr)}$ is obtained from the consistency equation as:

$$\lambda^{(corr)} = \frac{F}{\left(\frac{\partial F}{\partial\boldsymbol{\sigma}}\right)^T\mathbf{D}^{el}\frac{\partial Q}{\partial\boldsymbol{\sigma}} - \frac{\partial F}{\partial p_o^*}\frac{\partial p_o^*}{\partial\boldsymbol{\varepsilon}^{pl}}\mathbf{m}^T\frac{\partial Q}{\partial\boldsymbol{\sigma}}} \tag{37}$$

where F is the (positive) value of the yield function corresponding to the uncorrected stress state lying outside the yield locus. The correction of plastic strain is computed from (10), and the correction of hardening parameter from (11). The proposed drift correction scheme is described in more details in Sołowski and Gallipoli [22]. In the unlikely situation when, after the first correction, the stress state still lies outside the yield locus, the above algorithm is used again until the stress state is mapped back within the set tolerance.

The derivatives and stiffness matrix appearing in Eqs. (36) and (37) may be calculated either at the initial stress state, i.e. at the start of the substep, corresponding to $\boldsymbol{\sigma}$, s and p_0^* or at the updated but uncorrected stress state, i.e. at the end of the substep, corresponding to $\boldsymbol{\sigma} + \delta\boldsymbol{\sigma}$, $s + \delta s$ and $p_0^* + \delta p_0^*$. Although there are valid reasons for calculating these functions at the initial stress state, see e.g. Potts and Zdravkovic [11], the authors have found that, for lenient integration tolerances (i.e. relatively large values of SSTOL) and for higher order Runge–Kutta schemes (so that the substep size is particularly large), such functions must be evaluated at the updated stress state in order to ensure convergence. In the computations presented in the companion paper [24], the drift correction algorithm was always employed by evaluating the relevant derivatives and stiffness matrix at the updated stress state. This proved to be generally not worse than evaluating such functions at the initial stress state and substantially better in some cases. Additional supporting evidence in this respect may be found in Sołowski and Gallipoli [22].

After drift correction, a further check on the relative integration error R should be performed by adding the drift correction to the estimated integration error from Eqs. (26) and (27). This is

necessary to avoid large and improper drift corrections, especially when lenient values of SSTOL are used.

9. Extrapolation method

The main motivation for using the EPUS method is that it allows greater accuracy with respect to the EPS method in controlling the maximum integration error. However, during extensive testing of the previous Runge–Kutta schemes with different sets of BBM parameter values, the EPUS method has often proven inefficient and unstable. This was observed, for example, when the integration algorithm was used for the calibration of BBM by fitting computed results to laboratory tests. The unsatisfactory response of the EPUS method is probably due to two reasons, i.e. the failure of the error estimation when the stress substep at the denominator of Eqs. (33) and (34) is close to zero leading to an undetermined ratio and the numerical error due to limited machine accuracy when a large number of very small substeps are calculated. Although the latter effect can be partly countered by introducing extra accurate storage so that numerical values are recorded with doubled number of digits compared to the standard 32 bit format, machine accuracy will inevitably play an important role when computation of a large number of very small substeps is required. It is also important to consider that any potential benefit arising from the integration of a large number of small substeps is partially offset by the longer computation time.

Integration algorithms based on extrapolation methods have therefore been explored as an alternative to Runge–Kutta schemes, with the objective of verifying whether these algorithms could be used in combination with the EPUS method without affecting computational stability.

In the extrapolation method an initial approximation of the net stress increment $\Delta\boldsymbol{\sigma}_0^{(0)}$ is calculated by dividing the corresponding increment of strains and suction $\Delta\boldsymbol{\varepsilon}^{enh}$ in a number N_0 of equal subincrements $\delta\boldsymbol{\varepsilon}^{enh}$ (i.e. $\delta\boldsymbol{\varepsilon}^{enh} = \Delta\boldsymbol{\varepsilon}^{enh}/N_0$) and by using an explicit scheme to integrate the net stress over each of these subincrements. Subsequently, a second approximation of the net stress increment $\Delta\boldsymbol{\sigma}_1^{(0)}$ over the same increment $\Delta\boldsymbol{\varepsilon}^{enh}$ is calculated by using a number of subincrements $N_1 > N_0$. The linear combination of these two approximations yields a more accurate extrapolated approximation $\Delta\boldsymbol{\sigma}_1^{(1)}$ (i.e. an approximation whose error series has a leading term of higher order than $\Delta\boldsymbol{\sigma}_0^{(0)}$ and $\Delta\boldsymbol{\sigma}_1^{(0)}$). The error estimate of such extrapolated approximation is then computed and, if such error is acceptable, the calculations are terminated. Otherwise another approximation of the net stress increment $\Delta\boldsymbol{\sigma}_2^{(0)}$ is calculated with a number of subincrements $N_2 > N_1$. Similarly as above, the linear combination of $\Delta\boldsymbol{\sigma}_2^{(0)}$ and $\Delta\boldsymbol{\sigma}_1^{(0)}$ yields the extrapolated approximation $\Delta\boldsymbol{\sigma}_2^{(1)}$, which is then combined again with the previous extrapolated approximation $\Delta\boldsymbol{\sigma}_1^{(1)}$ to yield an even more accurate extrapolated approximation $\Delta\boldsymbol{\sigma}_2^{(2)}$. If the error estimate is acceptable at this stage, calculations are terminated. Otherwise a new approximation of the net stress increment $\Delta\boldsymbol{\sigma}_i^{(0)}$ is calculated by using a larger number of subincrements $N_i > N_{i-1}$ and this approximation is used for calculating further extrapolated values. Such process continues until the error estimate becomes smaller than the set tolerance.

The scheme is summarized in the following table:

$$\begin{array}{cccccccc}
N_0 & \Delta\boldsymbol{\sigma}_0^{(0)} & & & & & & \\
N_1 & \Delta\boldsymbol{\sigma}_1^{(0)} & \Delta\boldsymbol{\sigma}_1^{(1)} & & & & & \\
N_2 & \Delta\boldsymbol{\sigma}_2^{(0)} & \Delta\boldsymbol{\sigma}_2^{(1)} & \Delta\boldsymbol{\sigma}_2^{(2)} & & & & \\
N_3 & \Delta\boldsymbol{\sigma}_3^{(0)} & \Delta\boldsymbol{\sigma}_3^{(1)} & \Delta\boldsymbol{\sigma}_3^{(2)} & \Delta\boldsymbol{\sigma}_3^{(3)} & & & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & & \\
\Delta\boldsymbol{\sigma}_i^{(0)} & \Delta\boldsymbol{\sigma}_i^{(1)} & \Delta\boldsymbol{\sigma}_i^{(2)} & \ldots & \Delta\boldsymbol{\sigma}_i^{(3)} & \ldots & \Delta\boldsymbol{\sigma}_i^{(m)}
\end{array} \tag{38}$$

The first column of net stress increments in table (38) show the non-extrapolated approximation $\Delta\boldsymbol{\sigma}_i^{(0)}$ calculated by using an explicit integration scheme over N_i equal sized subincrements $\delta\boldsymbol{\varepsilon}_i^{enh}$ (i.e. $\delta\boldsymbol{\varepsilon}_i^{enh} = \Delta\boldsymbol{\varepsilon}^{enh}/N_i$). The other columns of table (38) contain the extrapolated approximation $\delta\boldsymbol{\sigma}_i^{(m)}$ (with $m > 0$), which can be calculated as a linear combination of the approximations $\Delta\boldsymbol{\sigma}_i^{(m-1)}$ and $\Delta\boldsymbol{\sigma}_{i-1}^{(m-1)}$ according to the following two alternative rules

$$\Delta\boldsymbol{\sigma}_i^{(m)} = \Delta\boldsymbol{\sigma}_i^{(m-1)} + \frac{\Delta\boldsymbol{\sigma}_i^{(m-1)} - \Delta\boldsymbol{\sigma}_{i-1}^{(m-1)}}{\left(\frac{N_i}{N_{i-m}}\right)^2 - 1} \tag{39}$$

or

$$\Delta\boldsymbol{\sigma}_i^{(m)} = \Delta\boldsymbol{\sigma}_i^{(m-1)} + \frac{\Delta\boldsymbol{\sigma}_i^{(m-1)} - \Delta\boldsymbol{\sigma}_{i-1}^{(m-1)}}{\left(\frac{N_i}{N_{i-m}}\right) - 1} \tag{40}$$

where $i,m = 1,2,3\ldots$

It can be shown that Eqs. (39) and (40) provide extrapolated approximations whose error expansion series have leading terms of progressively higher order and therefore converge towards the true solution (see, for example, Gragg [7]). The extrapolation rule of Eq. (39) is used when the error expansion series of the non-extrapolated initial approximations $\Delta\boldsymbol{\sigma}_i^{(0)}$ contains only even-powered terms such as:

$$\Delta\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma}_i^{(0)} + a_1\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^2 + a_2\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^4 + a_3\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^6 + \ldots \tag{41}$$

Eq. (40) is instead used when the error expansion series of the non-extrapolated initial approximations $\Delta\boldsymbol{\sigma}_i^{(0)}$ contains both odd and even-powered terms such as:

$$\Delta\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma}_i^{(0)} + a_1\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^2 + a_2\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^3 + a_3\left(\frac{\Delta\boldsymbol{\varepsilon}^{enh}}{N_i}\right)^4 + \ldots \tag{42}$$

If the error expansion of the approximations $\Delta\boldsymbol{\sigma}_i^{(0)}$ has the form of Eq. (41), (and Eq. (39) can therefore be used for extrapolation) the algorithm converges faster, as each subsequent extrapolation increases by two the order of the leading term in the error series. Eq. (39) is derived from consecutive eliminations of progressively higher order error terms in the expansion series of the net stress increment $\Delta\boldsymbol{\sigma}$. This is achieved by linear combination of subsequent couples of extrapolated series with leading error terms of the same order, starting from the linear combination of two series given by Eq. (41) with N_i and N_{i-1} subincrements, respectively. In the same way, Eq. (40) is derived by consecutive eliminations of progressively higher order error terms via linear combination of extrapolated series of the net stress increment $\Delta\boldsymbol{\sigma}$, starting from the linear combination of two series given by Eq. (42).

Note that the superscript $^{(m)}$ in the approximation $\delta\boldsymbol{\sigma}_i^{(m)}$ represents the number of subsequent extrapolations used to calculate that particular approximation. Thus those approximations with the same superscript $^{(m)}$ have leading terms of the error series of the same order (i.e. those approximations that are on the same column of table (38)). The subscript $_i$ in the approximation $\delta\boldsymbol{\sigma}_i^{(m)}$ identifies instead those approximations that belong to the same sequence of extrapolations with the first approximation corresponding to the explicit integration over N_i subincrements (i.e. those approximations that are on the same row of table (38)).The sequence of subincrements numbers N_i in Table (38) is subjective as it only needs to satisfy the requirement that subincrements numbers must be even and increasing. The most common sequences are:

$$N_i = 2(i+1) \Rightarrow N = \{2,4,6,8,10,12,14,\ldots\} \tag{43}$$

(Deuflhard [4])

$$N_i = 2N_{i-2} \Rightarrow N = \{2,4,6,8,12,16,24,32,48,64,\ldots\} \quad (44)$$

(Stoer and Bulirsch [25], Lambert [8])

$$N_i = 2^i \Rightarrow N = \{2,4,8,16,32,64,128,\ldots\} \quad (45)$$

(Lambert [8])

In this paper the following sequence was used:

$$N = \{32,48,64,96,128,160,192,256,320,384,448,512,608,736,992\} \quad (46)$$

This sequence performed best for the fairly large strain increments $\Delta\boldsymbol{\varepsilon}^{enh}$ used in this work, although for smaller increments the sequences given in (43), and (44) are advantageous. The chosen sequence of subincrements may seriously influence the quality of the results and the computational time required for integration.

Extrapolation methods have been previously used by applied mathematicians for the integration of ordinary differential equation, although their application to the integration of elasto-plastic soil models is unknown to the authors. They originate from the ideas of Richardson extrapolation (see Gragg [7]) and the particular extrapolation method used in this work is similar to the approach described in Stoer and Bulirsch [25]. The preferred explicit scheme for integrating the non-extrapolated approximations $\Delta\boldsymbol{\sigma}_i^{(0)}$ is the modified midpoint method (Deuflhard and Bornemann [4]) because the error series contains only even-powered terms as shown in Eq. (41). This makes the algorithm converge faster as consecutive extrapolations will increase the order of the leading error term by two. In this work, however, due to convergence problems of the modified midpoint method, the explicit integration of the non-extrapolated approximations $\Delta\boldsymbol{\sigma}_i^{(0)}$ has been performed by using the Runge–Kutta midpoint method. In this case, the error

Table A1
Modified Euler – Runge–Kutta 2(1) (Sloan [19]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	$\frac{1}{2}$	1
2	1	1	$\frac{1}{2}$	0

Table A2
Nystrom – Runge–Kutta 3(2) (Lee and Schiesser [9]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	–	$\frac{2}{8}$	$\frac{1}{4}$
2	$\frac{2}{3}$	$\frac{2}{3}$	–	$\frac{3}{8}$	$\frac{3}{4}$
3	$\frac{2}{3}$	0	$\frac{2}{3}$	$\frac{3}{8}$	0

Table A3
Bogacki–Shampine – Runge–Kutta 3(2) FSAL "first same as last" procedure (Bogacki and Shampine [2] – www.netlib.org/ode/rksuite).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	–	–	$\frac{2}{9}$	$\frac{7}{24}$
2	$\frac{1}{2}$	$\frac{1}{2}$	–	–	$\frac{1}{3}$	$\frac{1}{4}$
3	$\frac{3}{4}$	0	$\frac{3}{4}$	–	$\frac{4}{9}$	$\frac{1}{3}$
4	1	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{4}{9}$	0	$\frac{1}{8}$

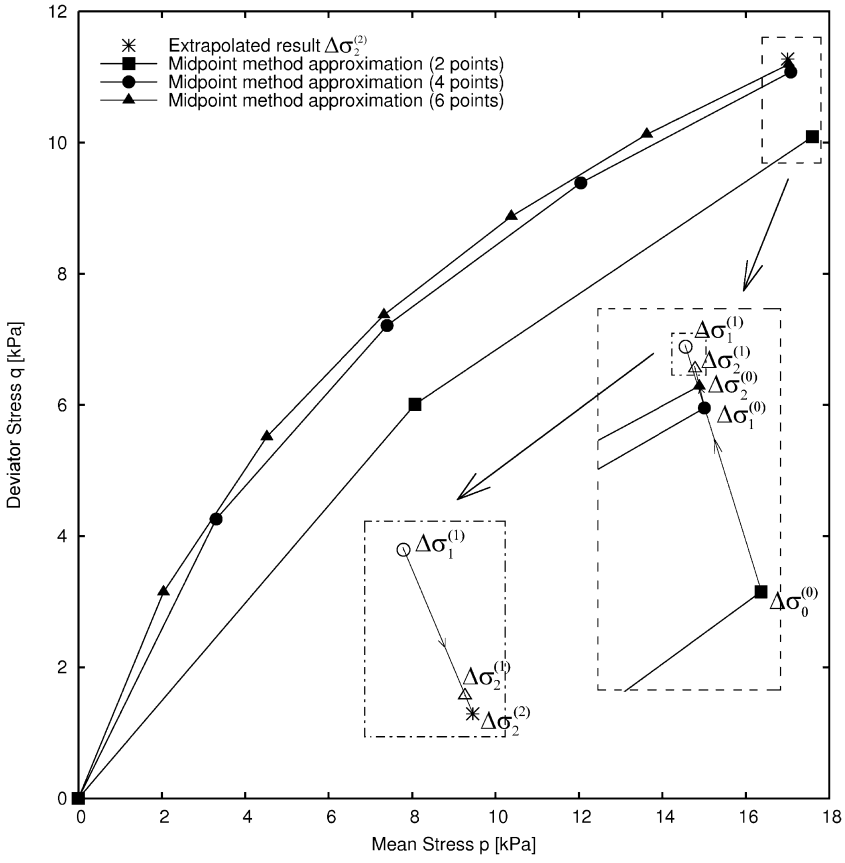

Fig. 1. Extrapolation method: stress increments $\Delta\boldsymbol{\sigma}_0^{(0)}$, $\Delta\boldsymbol{\sigma}_1^{(0)}$ and $\Delta\boldsymbol{\sigma}_2^{(0)}$ are first explicitly integrated from the same increment of strains and suction over 2, 4 and 6 subincrements, respectively. Subsequently these three stress increments are extrapolated to obtain the two stress increments $\Delta\boldsymbol{\sigma}_1^{(1)}$ and $\Delta\boldsymbol{\sigma}_2^{(1)}$, which are further extrapolated to obtain the most accurate stress increment $\Delta\boldsymbol{\sigma}_2^{(2)}$.

Table A4
Runge–Kutta 4(3). Error estimate coefficients $e^{(j)}$ given instead of coefficients $d^{(j)}$ (see Eq. (26)) (Lee and Schiesser [9]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$a^{(j4)}$	$b^{(j)}$	$e^{(j)}$
1	0	–	–	–	–	$\frac{1}{10}$	$-\frac{1}{15}$
2	$\frac{1}{3}$	$\frac{1}{3}$	–	–	–	0	0
3	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	–	–	$\frac{3}{10}$	$\frac{3}{10}$
4	$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{3}{8}$	–	$\frac{4}{10}$	$-\frac{4}{15}$
5	1	$\frac{1}{2}$	0	$-\frac{3}{2}$	2	$\frac{2}{10}$	$\frac{1}{30}$

Table A5
England – Runge–Kutta 5(4). Error estimate coefficients $e^{(j)}$ given instead of coefficients $d^{(j)}$ (see Eq. (26)) (Sloan [19], Lee and Schiesser [9]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$a^{(j4)}$	$a^{(j5)}$	$b^{(j)}$	$e^{(j)}$
1	0	–	–	–	–	–	$\frac{14}{336}$	$-\frac{42}{336}$
2	$\frac{1}{2}$	$\frac{1}{2}$	–	–	–	–	0	0
3	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	–	–	–	0	$-\frac{224}{336}$
4	1	0	−1	2	–	–	$\frac{35}{336}$	$-\frac{21}{336}$
5	$\frac{2}{3}$	$\frac{7}{27}$	$\frac{10}{27}$	0	$\frac{1}{27}$	–	$\frac{162}{336}$	$\frac{162}{336}$
6	$\frac{2}{10}$	$\frac{28}{625}$	$-\frac{125}{625}$	$\frac{546}{625}$	$\frac{54}{625}$	$-\frac{378}{625}$	$\frac{125}{336}$	$\frac{125}{336}$

Table A6
Cash Karp – Runge–Kutta 5(4) (Press et al. [12], Lee and Schiesser [9]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$a^{(j4)}$	$a^{(j5)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	–	–	–	–	$\frac{37}{378}$	$\frac{2825}{27648}$
2	$\frac{2}{10}$	$\frac{2}{10}$	–	–	–	–	0	0
3	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$	–	–	–	$\frac{250}{621}$	$\frac{18575}{48384}$
4	$\frac{6}{10}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{12}{10}$	–	–	$\frac{125}{594}$	$\frac{13525}{55296}$
5	1	$-\frac{11}{54}$	$\frac{5}{2}$	$-\frac{70}{27}$	$\frac{35}{27}$	–	0	$\frac{277}{14336}$
6	$\frac{7}{8}$	$\frac{1631}{55296}$	$\frac{175}{512}$	$\frac{575}{13824}$	$\frac{44275}{110592}$	$\frac{253}{4096}$	$\frac{512}{1771}$	$\frac{1}{4}$

expansion only approximates that of Eq. (41) (see Deuflhard and Bornemann [4] for a proof) and thus the extrapolation formula of Eq. (39), based on the error expansion of Eq. (41), is not rigorously correct. The explicit integration with the Runge–Kutta midpoint method is based on the use of the tangent elasto-plastic matrix, $\mathbf{D}^{ep}$ and hardening matrix, $\mathbf{H}$ defined in Eqs. (14) and (18), respectively. Details about the Runge–Kutta midpoint method are provided in Sołowski and Gallipoli [22].

The extrapolation method does not have a fixed order, as Runge–Kutta schemes, but the order of the approximation depends on the number of subsequent extrapolations, i.e. on the number of columns in Eq. (38). Each subsequent column has another leading term eliminated in the error series and thus gives a higher order approximation. The idea of getting an improved extrapolated result from the linear combination of two previous less accurate solutions is graphically illustrated in Fig. 1. Three approximated net stress increments $\Delta\boldsymbol{\sigma}_0^{(0)}$, $\Delta\boldsymbol{\sigma}_1^{(0)}$ and $\Delta\boldsymbol{\sigma}_2^{(0)}$ are computed from the same increment of strains and suction by using the Runge–Kutta midpoint method over two, four and six subincrements, respectively. From these three stress increments, the first two extrapolated approximations, $\Delta\boldsymbol{\sigma}_1^{(1)}$ and $\Delta\boldsymbol{\sigma}_2^{(1)}$, are obtained according to Eq. (39). These two extrapolated approximations are subsequently used again in Eq. (39) to obtain a higher order extrapolated result, $\Delta\boldsymbol{\sigma}_2^{(2)}$, which is more accurate than any previous solution.

The estimate $\mathbf{E}(\Delta\boldsymbol{\sigma})$ of the integration error over the strain $\Delta\boldsymbol{\varepsilon}^{enh}$ is obtained as (see Deuflhard and Bornemann [4], Stoer and Bulirsch [25]):

$$\mathbf{E}(\Delta\boldsymbol{\sigma}) = \Delta\boldsymbol{\sigma}_i^{(m)} - \Delta\boldsymbol{\sigma}_i^{(m-1)} \tag{47}$$

The suitability of this estimated error is then assessed by using either EPS or EPUS control against a set tolerance as previously described. The robustness of the extrapolation method, even when used with the less stable EPUS control, is noticeably greater than any of the adaptive Runge–Kutta methods. This is partly explained by the fact that the extrapolation method provides an estimate of the global error (i.e. the error over the whole increment) as opposed to the adaptive Runge–Kutta schemes, which provide an estimate of the local error (i.e. the error in each substep). Further information about the extrapolation method is provided in Sołowski [21] and Sołowski and Gallipoli [23].

Table A7
Dormand Prince – Runge–Kutta 5(4) FSAL "first same as last" procedure (Dormand [5]).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$a^{(j4)}$	$a^{(j5)}$	$a^{(j6)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	–	–	–	–	–	$\frac{35}{384}$	$\frac{5179}{57600}$
2	$\frac{2}{10}$	$\frac{2}{10}$	–	–	–	–	–	0	0
3	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$	–	–	–	–	$\frac{500}{1113}$	$\frac{7571}{16695}$
4	$\frac{8}{10}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$	–	–	–	$\frac{125}{192}$	$\frac{393}{640}$
5	$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2178}$	$\frac{64448}{6561}$	$-\frac{212}{729}$	–	–	$-\frac{2187}{6784}$	$-\frac{92097}{339200}$
6	1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$	–	$\frac{11}{84}$	$\frac{187}{2100}$
7	1	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2178}{6784}$	$\frac{11}{84}$	0	$\frac{1}{40}$

Table A8
Bogacki–Shampine – Runge–Kutta 5(4) FSAL "first same as last" procedure (Bogacki and Shampine [2] – www.netlib.org/ode/rksuite).

Stage number	$c^{(j)}$	$a^{(j1)}$	$a^{(j2)}$	$a^{(j3)}$	$a^{(j4)}$	$a^{(j5)}$	$a^{(j6)}$	$a^{(j7)}$	$b^{(j)}$	$d^{(j)}$
1	0	–	–	–	–	–	–	–	$\frac{578}{8064}$	$\frac{2479}{34992}$
2	$\frac{1}{6}$	$\frac{1}{6}$	–	–	–	–	–	–	0	0
3	$\frac{2}{9}$	$\frac{2}{27}$	$\frac{4}{27}$	–	–	–	–	–	$\frac{4440339}{15491840}$	$\frac{123}{416}$
4	$\frac{3}{7}$	$\frac{183}{1372}$	$-\frac{162}{343}$	$\frac{1053}{1372}$	–	–	–	–	$\frac{24353}{124800}$	$\frac{612941}{3411720}$
5	$\frac{2}{3}$	$\frac{68}{297}$	$-\frac{4}{11}$	$\frac{42}{143}$	$\frac{1960}{3861}$	–	–	–	$\frac{387}{44800}$	$\frac{43}{1440}$
6	$\frac{3}{4}$	$\frac{597}{22528}$	$\frac{81}{352}$	$\frac{63099}{585728}$	$\frac{58653}{366080}$	$\frac{4617}{20480}$	–	–	$\frac{2152}{5985}$	$\frac{2272}{6561}$
7	1	$\frac{174197}{959244}$	$-\frac{30942}{79937}$	$\frac{8152137}{19744439}$	$\frac{666106}{1039181}$	$-\frac{29421}{29068}$	$\frac{482048}{414219}$	–	$\frac{7267}{94080}$	$\frac{79937}{1113912}$
8	1	$\frac{587}{8064}$	0	$\frac{4440339}{15491840}$	$\frac{24353}{124800}$	$\frac{387}{44800}$	$\frac{2152}{5985}$	$\frac{7267}{94080}$	0	$\frac{3293}{556956}$

Note that this Bogacki–Shampine pair has an additional error estimate $e^{(j)'}$, obtained from lower order solution different from the one given in the above table. This error estimate has the following coefficients $\left[-\frac{3}{1280}, 0, \frac{6561}{632320}, -\frac{343}{20800}, \frac{243}{12800}, -\frac{1}{95}, 0\right]$ and does not require calculation of the last stage of the procedure. In case such error estimate is too large, the last stage of the procedure is skipped and the substep is rejected.

10. Conclusions

The paper describes the application of Runge–Kutta schemes of different order, as well as a novel extrapolation scheme, to the explicit stress integration of an elasto-plastic unsaturated soil model with automatic error control. The extrapolation scheme, based on the idea of Richardson extrapolation, has not been previously used for stress integration of elasto-plastic soil models and presents significant advantages in terms of robustness and stability with respect to conventional adaptive Runge–Kutta integration.

Two different methods are described to enforce automatic error control during stress integration, i.e. the Error Per Step (EPS) and the Error Per Unit Step (EPUS). The former method controls the error relative to the current magnitude of net stresses and has been conventionally used in adaptive stress integration. The latter method controls the error relative to the corresponding change of net stress and, although it is preferable in terms of accuracy, has been seldom implemented in explicit adaptive integration due to lack of stability and large computational cost.

The developments presented in this paper refer specifically to the Barcelona Basic Model, however they do not address possible non-convexity of the loading collapse (LC) curve and neglect yielding on the suction increase (SI) line. The framework of the proposed algorithms might also be applicable to other elasto-plastic models for unsaturated soils formulated in terms of two independent stress variables (such as net stress and suction).

An extensive programme of numerical tests is presented in a companion paper [24] to compare the efficiency, accuracy and robustness of the various integration schemes here described. This also includes a comparison of the results from the different integration schemes against analytical solutions to assess the accuracy of each scheme and the performance of the chosen error control method.

Acknowledgements

The authors gratefully acknowledge funding by the European Commission through the MUSE Research Training Network, contract: MRTN-CT-2004-506861.

Appendix A. Runge–Kutta coefficient

This appendix provides the values of the coefficients $a^{(jk)}$, $b^{(j)}$, $c^{(j)}$ and $d^{(j)}$, as defined in Eqs. (23)–(26), for the different Runge–Kutta schemes considered in this work. These coefficients have also been used in the numerical tests presented in the companion paper [24] (see Tables A1–A8).

References

[1] Alonso EE, Gens A, Josa A. A constitutive model for partially saturated soils. Géotechnique 1990;40(3):405–30.
[2] Bogacki P, Shampine LF. An efficient Runge–Kutta (4,5) pair. Computers Math Appl 1996;32(6):15–28 [Method coefficients are given in www.netlib.org/ode/rksuite (last accessed August 2008)].
[3] Borja RI. Cam-clay plasticity. Part V: a mathematical framework for three-phase deformation and strain localization analyses of partially saturated porous media. Comput Meth Appl Mech Eng 2004;193(48–51):5301–38.
[4] Deuflhard P, Bornemann F. Scientific computing with ordinary differential equations. Text in applied mathematics 42. Springer; 2002.
[5] Dormand JR. Numerical methods for differential equations. A computational approach. Boca Raton CRC Press; 1996.
[6] Gear WG. Numerical initial value problems in ordinary differential equations. Englewood Cliffs: Prentice Hall Inc.; 1971.
[7] Gragg WB. On extrapolation algorithms for ordinary initial value problems. SIAM J Numer Anal 1965;2:384–403.
[8] Lambert JB. Computational methods in ordinary differential equations. John Wiley & Sons; 1973.
[9] Lee HJ, Schiesser WE. Ordinary and partial differential equation routines in C, C++, Fortran, Java, Maple, and Matlab. Boca Raton: CRC Press; 2003.
[10] Potts DM, Gens A. A critical assessment of methods of correcting for drift from the yield surface in elasto-plastic finite elements analysis. Int J Numer Anal Meth Geomech 1985;9:149–59.
[11] Potts DM, Zdravkovic L. Finite element analysis in geotechnical engineering. Theory. London: Thomas Telford Publishing; 1999.
[12] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C++. The art of scientific computing. 2nd ed. Cambridge University Press; 2002.
[13] Shampine LF. Numerical solution of ordinary differential equations. New York: Chapman & Hall; 1994.
[14] Sheng D, Gens A, Fredlund DG, Sloan SW. Unsaturated soils: from constitutive modelling to numerical algorithms. Comput Geotech 2008;35:810–24.
[15] Sheng D, Sloan SW, Gens A, Smith DW. Finite element formulation and algorithms for unsaturated soils. Part I: theory. Int J Numer Anal Meth Geomech 2003;27:745–65.
[16] Sheng D, Sloan SW, Gens A, Smith DW. Finite element formulation and algorithms for unsaturated soils. Part II: verification and application. Int J Numer Anal Meth Geomech 2003;27:767–90.
[17] Sheng D, Sloan SW, Yu HS. Aspects of finite element implementation of critical state models. Comput Mech 2000;26:185–96.
[18] Simo JC, Hughes TJR. Computational inelasticity. Interdisciplinary applied mathematics series, vol. 7. Springer; 1998.
[19] Sloan SW. Substepping schemes for the numerical integration of elastoplastic stress–strain relations. Int J Numer Meth Eng 1987;24:893–911.
[20] Sloan SW, Abbo AJ, Sheng D. Refined explicit integration of elastoplastic models with automatic error control. Eng Comput 2001;18(1/2):121–54.
[21] Sołowski WT. Unsaturated soils: constitutive modelling and explicit stress integration. PhD Thesis, Durham University; 2008.
[22] Sołowski WT, Gallipoli D. A stress–strain integration algorithm for unsaturated soil elastoplasticity with automatic error control. In: Schweiger HF, editor. Numerical methods in geotechnical engineering – proceedings of the 6th European conference on numerical methods in geotechnical engineering, Graz, Austria, 6–8 September 2006. Taylor & Francis Group; 2006. p. 113–9.
[23] Sołowski WT, Gallipoli D. Numerical integration of elasto-plastic constitutive models using the extrapolation method. In: Pande S, Pietruszczak GN, editors. Proceeding of the NUMOG X conference, Rhodes, Greece, 25–27 April 2007. Taylor & Francis Group; 2007. p. 211–7.
[24] Sołowski WT, Gallipoli D. Explicit stress integration with error control for the Barcelona Basic Model. Part II: algorithms efficiency and accuracy. Comput Geotech 2010;37(1/2):68–81.
[25] Stoer J, Bulirsch R. Introduction to numerical analysis. 3rd ed. Springer; 2002.
[26] Vaunat J, Cante JC, Ledesma A, Gens A. A stress point algorithm for an elastoplastic model in unsaturated soils. Int J Plasticity 2000;16:121–41.
[27] Wheeler SJ, Gallipoli D, Karstunen M. Comments on use of the Barcelona Basic Model for unsaturated soils. Int J Numer Anal Meth Geomech 2002;26:1561–71.
[28] Zhang HW, Heeres OM, de Borst R, Schrefler BA. Implicit integration of a generalized plasticity constitutive model for partially saturated soil. Eng Comput 2001;18(1/2):314–36.
[29] Zhao J, Sheng D, Rouainia M, Sloan SW. Explicit stress integration of complex soil models. Int J Numer Anal Meth Geomech 2005;29:1209–29.