An r-h Adaptive Kinematic Approach for 3D Limit Analysis

Zhenhao Shi[a], James P. Hambleton[a,b,*]

[a]Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208, USA
[b]School of Engineering, University of Newcastle, Callaghan, NSW 2308, Australia

Abstract

This paper explores a pathway for increasing efficiency in numerical 3D limit analysis through r-h adaptivity, wherein nodal positions (r) and element lengths (h) are successively refined. The approach uses an iterative, nested optimization procedure involving three components: (1) determination of velocities for a fixed mesh of rigid, translational elements (blocks) using second-order cone programming; (2) adaptation of nodal positions using non-linear optimization (r adaptivity); and (3) subdivision of elements based on the magnitude of the velocity jumps (h adaptivity). Examples show that the method can compute reasonably accurate limit loads at relatively low computational cost.

Keywords: limit analysis, 3D, kinematic method, upper bound, adaptivity

*Corresponding author. Address: Department of Civil and Environmental Engineering, Northwestern University, Technological Institute A122, 2145 Sheridan Road, Evanston, IL 60208; Tel.: +1 847 491 4858

Email addresses: zhenhao.shi@northwestern.edu (Zhenhao Shi), jphambleton@northwestern.edu (James P. Hambleton)

1. Introduction

Accurate evaluation of the limit load, or collapse load, causing failure of a mass of geomaterial is crucial for the design of geotechnical infrastructure, including foundations, slopes, and earth retaining systems. Limit load computations are also central in the determination of how to induce failure deliberately, as in excavation, mining, and earthmoving (e.g., Hettiaratchi and Reece, 1974; Godwin and O'Dogherty, 2007; Hambleton et al., 2014; Hambleton, 2017). Many models rely on a two-dimensional (2D) idealization of the true configuration (e.g., plane strain or axisymmetry), which significantly simplifies the calculations. However, in many cases, the three-dimensional (3D) nature of the problem cannot be ignored. When 3D conditions prevail, computations based on the 2D simplification can overestimate or underestimate the limit load (Soubra and Regenass, 2000; Antão et al., 2011; Michalowski, 2001; Griffiths and Marquez, 2007; Michalowski and Drescher, 2009; Wörden and Achmus, 2013).

Among various existing methods, the kinematic approach of limit analysis is a particularly effective and useful means of evaluating limit loads (cf. Chen, 1975). The kinematic theorem states that for any kinematically admissible velocity field (i.e., failure or collapse mechanism), the load computed by equating the work rate of external forces to the internal energy dissipation rate is a rigorous bound on the true limit load. It gives an upper bound for a load inducing collapse and a lower bound for a load resisting collapse (Drescher, 1991). A kinematically admissible velocity field is one that satisfies boundary conditions and the plastic flow rule. The kinematic theorem requires that material is perfectly plastic and obeys the associative flow rule. The consequences of associativity and possible workarounds in instances where it may lead to unrealistic predictions are discussed by various authors (Davis, 1968; Davis and Booker, 1971; Chen, 1975; Drescher and Detournay, 1993; Krabbenhoft et al., 2012; Sloan, 2013). For 3D problems with simple geometries and loading conditions, a kinematically admissible mechanism can be constructed manually, thereby permitting an analytical or semi-analytical solution (e.g., Murray and Geddes 1987; Soubra and Regenass 2000; Michalowski 2001). Nevertheless, it is generally difficult to construct collapse mechanisms for 3D problems, and numerical methods are usually necessary.

Finite element limit analysis (FELA) is a powerful numerical implementation that can evaluate 3D collapse loads without assuming a failure mechanism *a priori* (Lyamin and Sloan, 2002a,b; Lyamin et al., 2007; Vicente da Silva and Antão, 2008; Krabbenhøft et al., 2008; Martin and Makrodimopoulos, 2008; Sloan, 2013). As in the conventional finite element method (FEM), FELA discretizes the domain into elements and interpolates the velocity field based on discrete values at nodes and the assumed shape functions. The optimal velocity field is computed by solving a large-scale optimization problem. The objective function corresponds to the limit load, and the unknown nodal velocities are constrained by enforcing kinematically admissibility. In FELA, a certain discretization of the domain (i.e., meshing) leads only to a subset of all possible velocity fields. Therefore, the limit load computed by FELA is often highly sensitive to the finite element mesh, particularly in regions of localized deformation. To maximize the solution accuracy using a minimum number of elements, adaptive mesh refinement techniques (i.e., h adaptivity) have been proposed to automatically refine regions featuring large gradients (Borges et al., 1999, 2001; Lyamin et al., 2005; Martin, 2011) or large gaps between upper-bound and lower-bound solutions computed on the same mesh (Ciria et al., 2008; Muñoz et al., 2009). The concept of h adaptivity has played a key role in improving the accuracy and computational efficiency of 2D analyses. In contrast, 3D FELA based on adaptive mesh refinement (e.g., Dunne and Martin, 2017) and its performance have not been investigated in great detail.

Another general numerical approach referred to as discontinuity layout optimization (DLO) has been developed by Smith and Gilbert (2007) and Hawksbee et al. (2013) on the basis of optimizing a velocity field consisting only of so-called velocity discontinuities, which represent infinitesimally thin zones of shearing. DLO focuses on optimizing the arrangement of these discontinuities, with the tacit assumption that the material enclosed by the discontinuities is rigid. This method searches for an optimal combination of the possible discontinuities interconnecting a *fixed* grid of nodes. Because the grid is fixed, the grid resolution has to be refined to capture intricate features or reasonably represent a continuous velocity field, which can dramatically increase the number of potential discontinuities at the cost of computational expediency (Hawksbee et al., 2013).

While the above-mentioned numerical approaches represent valuable tools to evaluate limit loads for 3D problems, they tend to be computationally intensive. In many cases, the optimal mechanism is in fact relatively simple, and the standard formulations of FELA and DLO can be unnecessarily onerous. Furthermore, the computational demands of existing techniques impose a significant limitation for the emerging computational approach referred to in this paper as the sequential kinematic method (SKM). In SKM, kinematic solutions are sequentially computed as a means of simulating a full process of deformation, and the optimal velocity field within any particular increment is used to update the model geometry and material properties. Given its computational efficiency and stability, SKM has become a compelling alternative to conventional techniques such as FEM for simulating problems in which capturing the evolution of material boundaries is critical (Hambleton and Drescher, 2012; Mary et al., 2013; Hambleton et al., 2014; Kong et al., 2017). In particular, SKM shows a remarkable capability in modeling the large deformation of cohesionless soils (Hambleton et al., 2014; Kashizadeh et al., 2014), which poses a significant challenge for conventional approaches. Current SKM formulations, however, are restricted to 2D. Extension to 3D has been largely halted by the lack of efficient methods to compute the optimal velocity field within each increment of simulation.

In this work, we investigate the concept of r adaptivity, in combination with h adaptivity, and assess the potential of this approach for increasing computational efficiency in 3D limit analysis. Pioneered in the earlier work of Johnson (1995) and more recently explored for 2D limit analysis (Milani and Lourenço, 2009; Hambleton and Sloan, 2013; Milani, 2015; He and Gilbert, 2016; Muñoz et al., 2018), r adaptivity improves the computed limit load by explicitly optimizing the *nodal positions* that control the locations of possible velocity discontinuities. Because relatively coarse meshes with suitably placed edges (velocity discontinuities) are often sufficient to obtain accurate solutions, kinematic FELA and DLO enriched with r adaptivity offers a promising pathway for improving efficiency, as previously demonstrated for 2D formulations.

2. Overview of the r-h adaptive approach

The general concept we explore in this paper is to start with a simple velocity field, one requiring minimal computational effort, and then refine this field to improve the accuracy of the computed limit load and collapse mechanism. We adopt a formulation in which the velocity field is characterized by discrete regions (blocks or elements) of translational motion separated by velocity discontinuities. These elements are tetrahedral by assumption, such that the edges, representing velocities discontinuities, are planar. We restrict our attention to material obeying the Mohr-Coulomb yield criterion and assume that the internal friction angle ϕ and cohesion c are constant across the soil mass. Similarly, the material unit weight, denoted by γ, is assumed to be constant. Spatially varied ϕ, c, and γ can be included into the current formulation by constructing mesh according to the soil stratigraphy, in that no discontinuity spans across different layers of soils. In the case of inter-layer discontinuities, the highest angle of friction and cohesion encountered should be used to maintain the upper-bound status of the solution.

Starting from an initial arrangement (mesh) of elements, the proposed r-h adaptive approach proceeds iteratively, and each iteration involves three key components. First, as explained in detail in Section 3, the optimal velocities for a fixed mesh are determined using second-order cone programming (SOCP). Second, as described in Section 4, the nodal positions are regarded as variables determined through non-linear optimization (r adaptivity). Third, elements are potentially subdivided (h adaptivity). Section 5 and 6 explain this third step and how each of the three components are combined to obtain a complete solution algorithm, respectively. Section 7 considers several example problems to which the algorithm is applied.

3. Optimization of the velocity field for a fixed mesh

Considering an arbitrary mesh of rigid tetrahedral elements (blocks), Hambleton and Sloan (2016) proposed a technique that utilizes second-order cone programming (SOCP) to search for a kinematically admissible velocity field that yields an optimal limit load and collapse mechanism. For completeness, its mathematical formulation is summarized here.

A generic pair of blocks is depicted in Fig. 1(a). The velocity jump between these blocks is denoted by Δv_i and is calculated as $\Delta v_i = v_i^I - v_i^{II}$, where v_i^I and v_i^{II} are the block velocities. The superscripts I and II indicate, arbitrarily, the first and second block, and the index $i = 1, 2, 3$ indicates the

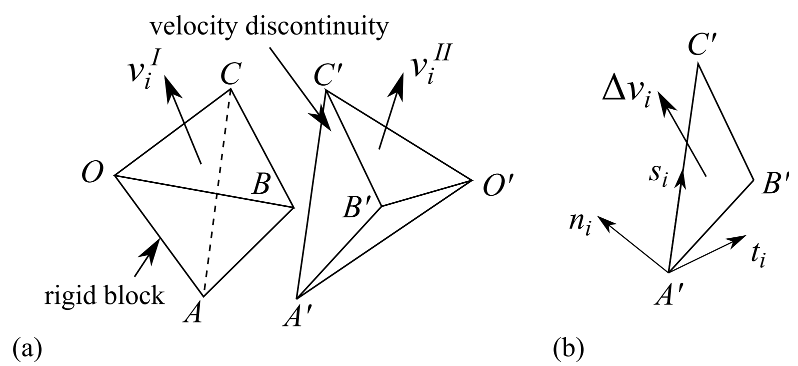


Figure 1: Schematics showing (a) 3D rigid blocks separated by a planar velocity discontinuity and (b) the definition of a local coordinate system associated with the discontinuity plane.

velocity component. In this work, the component associated with $i = 3$ is always in the vertical direction, and it is assumed to be positive when the velocity is upward (i.e., opposite the direction of gravity). Chen (1975) shows that for materials obeying the Mohr-Coulomb yield criterion, the energy dissipation rate along the planar velocity discontinuities between elements (blocks) can be expressed as

$$\dot{d} = cA|\Delta v_t| \tag{1}$$

The variable A denotes the area, and Δv_t is the tangential velocity jump with respect to the plane of the discontinuity. The absolute value is prescribed so that the dissipated power is always positive, regardless of the shearing direction. To fulfill the associative flow rule corresponding to the Mohr-Coulomb yield condition, a kinematically admissible velocity discontinuity has to meet the following "jump condition" (Chen, 1975):

$$\Delta v_n = |\Delta v_t| \tan\phi \tag{2}$$

The variable Δv_n denotes the normal velocity jump. By adopting the local coordinate system shown in Fig. 1(b), Eqs. (1) and (2) can be rewritten as

$$\begin{aligned} \dot{d} &= cA\sqrt{(\Delta v_i t_i)^2 + (\Delta v_i s_i)^2} \\ \Delta v_i n_i &= \tan\phi\sqrt{(\Delta v_i t_i)^2 + (\Delta v_i s_i)^2} \end{aligned} \tag{3}$$

Following the summation convention, the quantities $\Delta v_i t_i$ and $\Delta v_i s_i$ are dot products calculated, for example, as $\Delta v_i t_i = \Delta v_1 t_1 + \Delta v_2 t_2 + \Delta v_3 t_3$. In Eq. (3), n_i is a unit vector normal to the plane of the discontinuity, and t_i and s_i are two unit vectors parallel to the plane. These three vectors give a mutually orthogonal transformed basis for expressing the velocity vectors, as depicted in Fig. 1(b). Note that in accordance with measuring the velocity jump from block I to II discussed above, the vector n_i points towards block I such that a positive normal component of the velocity jump indicates dilation.

In order to write Eq. (3) in a form amenable to SOCP, the quantity $\sqrt{(\Delta v_i t_i)^2 + (\Delta v_i s_i)^2}$ is replaced by a dummy variable μ:

$$\begin{aligned} \dot{d} &= cA\mu \\ \Delta v_i n_i &= \mu \tan\phi \end{aligned} \tag{4}$$

The dummy variable μ is then constrained as follows:

$$\mu \geq \sqrt{(\Delta v_i t_i)^2 + (\Delta v_i s_i)^2} \tag{5}$$

Eq. (5) is in the form of a so-called second-order cone constraint, one of the types permitted in SOCP in addition to linear equality and inequality constraints (cf. Sturm, 2002).

We note that the expressions given by Eq. (4) are exact only in the particular instance where strict equality is achieved in Eq. (5):

$$\mu = \sqrt{(\Delta v_i t_i)^2 + (\Delta v_i s_i)^2} \tag{6}$$

Equality is achieved by constructing the optimization problem such that the dummy variable μ is minimized, and thus μ is driven to equality as in Eq. (6). Application to example problems, such as those considered in Section 7.3, reveals that equality is achieved in most cases. However, in the case of cohesionless material ($c = 0$) for which the dissipation $\dot{d}$ vanishes, equality is not always achieved. Nevertheless, it should be noted that, when the equality in Eq. (5) is not satisfied, the solution remains an upper bound of the true collapse load because the energy dissipation and the jump condition are effectively computed according to a larger cohesion and friction angle, respectively.

By equating the rate of internal energy dissipation to the work rate of external forces for an assembly of elements (blocks), one obtains

$$\sum_{j=1}^{N_D} \dot{d}_j = -\sum_{k=1}^{N_B} \gamma V_k v_{3k} + \int_{S^*} t_i^* v_i ds + \int_S t_i v_i ds \tag{7}$$

In Eq. (7), N_D and N_B are the number of discontinuity planes and the number of blocks, respectively, and subscripts j and k are used to indicate quantities corresponding to the j^{th} discontinuity plane and the k^{th} block. The variable V_k denotes the volume of the k^{th} block, a readily computed constant for a fixed mesh. The three terms on the right side of Eq. (7) represent the work rate of body forces, fixed surface tractions (t_i^*) and tractions along the surface where the limit loads is evaluated (t_i), respectively.

Drescher (1991), Sloan (1995), and Michalowski (2001) among others show how Eq. (7) can be manipulated to obtain various expressions of the limit load. A case encompassing all examples considered in Section 7 is that the direction of t_i is fixed, or known *a priori*, and velocities along the boundary S are uniform, as could occur for a rigid footing or translational retaining wall. In this instance, the unknown traction t_i is expressed in terms of a fixed traction t_i^* as $t_i = \lambda t_i^*$, where $\lambda \geq 0$ is an unknown multiplier dictating the magnitude of the limit load. The last term in Eq. (7) is $\int_S t_i v_i ds = v_i \int_S \lambda t_i^* ds = \lambda v_i F_i^*$, where $F_i^* = \int_S t_i^* ds$ is the resultant force. The magnitude of the velocity is arbitrary (cf. Chen, 1975), and thus one can write $v_i F_i^* = \alpha$, where α is an arbitrary constant. Equation (7) can then be manipulated to write:

$$\lambda = \frac{1}{\alpha}\left(\sum_{j=1}^{N_D} \dot{d}_j + \sum_{k=1}^{N_B} \gamma V_k v_{3k} - \int_{S^*} t_i^* v_i ds\right) \tag{8}$$

Here we assume α is unity (a value of 1 with appropriate units) for convenience. Depending on the distribution of the fixed tractions t_i^*, the final term in parenthesis in Eq. (7) can be integrated to obtain a sum over the unknown velocities, viz.

$$\int_{S^*} t_i^* v_i ds = \sum_{l=1}^{N_F} \beta_{il} v_{il} \tag{9}$$

In Eq. (9), N_F is the number of elements with fixed tractions, and β_{il} ($i = 1, 2, 3; l = 1, ..., N_F$) are constant coefficients. The notation v_{il} again indicates the i^{th} velocity component of the l^{th} element.

Finally, the optimization of the velocity field for a fixed mesh is written in the standard form of SOCP as follows:

$$\begin{aligned}
min \quad & \lambda = \sum_{j=1}^{N_D} \dot{d}_j + \sum_{k=1}^{N_B} \gamma V_k v_{3k} - \sum_{l=1}^{N_F} \beta_{il} v_{il} \\
s.t. \quad & \Delta v_{ij} n_{ij} = \mu_j \tan\phi \quad j = 1, ..., N_D \\
& \dot{d}_j = c A_j \mu_j \quad j = 1, ..., N_D \\
& \mu_j \geq \sqrt{(\Delta v_{ij} t_{ij})^2 + (\Delta v_{ij} s_{ij})^2} \quad j = 1, ..., N_D
\end{aligned} \tag{10}$$

For a load resisting collapse, where the work rate of the unknown tractions on the velocity is negative, the kinematic theorem of limit analysis leads to a lower bound on the true collapse load (cf. Drescher, 1991). To compute such a lower bound, Eq. (10) is converted to a maximization problem by minimizing the negative of the objective function. In this work, the Mosek toolbox integrated with MATLAB (Mosek, 2015) is employed to solve the SOCP problem.

Upon solving the SOCP problem of Eq. (10), one obtains an optimal value for the load multiplier, denoted by λ_{opt}. The computed bound on the true collapse load is then simply

$$F_i = \lambda_{opt} F_i^* \tag{11}$$

4. Optimization of nodal positions (r adaptivity)

The bound on the limit load computed using Eq. (10) depends strongly on the positions of the nodes within the mesh that define the locations of potential velocity discontinuities. In particular, the optimal velocity field and load multiplier λ_{opt} depend on the coordinates of the nodes that are not constrained by boundary conditions or symmetry, and are therefore free to move. The coordinates of these nodes are denoted by x_{im}. Index i again gives the component ($i = 1, 2, 3$), and index m ($i = 1, ..., N_R$) identifies each of the free nodes.

For the purpose of optimizing the nodal positions, a non-linear optimization problem is formulated as follows:

$$\begin{aligned} min \quad & \lambda_{opt}(x_{im}) \\ s.t. \quad & V_k(x_{im}) \geq 0 \quad k = 1, ..., N_B \\ & x_{im}^l \leq x_{im} \leq x_{im}^u \end{aligned} \tag{12}$$

This non-linear optimization is nested with the SOCP described above, in that the objective function in Eq. (12) is the load multiplier computed for a given set of nodal positions x_{im} ($i = 1, 2, 3$; $m = 1, ..., N_R$), defined and evaluated in precisely the same way as in Section 3. To prevent the interpenetration of elements and ensure computational stability, the first set of constraints in Eq. (12) requires that element volume V_k ($k = 1, ..., N_B$) is always positive. It should be noted that we permit the possibility $V_k = 0$, thus allowing elements to collapse to transition layers with zero thickness. The variables x^l_{im} and x^u_{im} appearing in the second set of inequality constraints define allowable limits for certain nodal position components. For instance, the z-coordinate of the ground surface is an upper bound on the position of all nodes along the z-direction.

Due to boundary conditions and symmetry, some of the position components (x, y, and z) are fixed. Rather than imposing constraints, the total number of free variables introduced in the non-linear optimization problem of Eq. (12) is condensed from $3N_R$ to DOF, where $DOF = 3N_R - N_{FC}$ and N_{FC} is the total number of fixed position components.

As the objective function and constraints are non-linear functions of the free (unknown) variables x_{im}, the optimization problem of Eq. (12) falls within the general domain of non-linear constrained optimization. A preliminary study employs two algorithms embedded in the FMINCON solver of MATLAB to solve this problem: the interior point method (IPM) and sequential quadratic programming (SQP). Both methods represent the state of the art in solving general constrained optimization problems. It is found that these two methods can achieve similar solutions. However, IPM requires more iterations, and during some iteration processes it diverges (i.e., the objective function increases rather than decreases). Accordingly, SQP is used throughout this work. It should be noted that the theoretical reason why

SQP outperforms IPM remains unclear. This is due in part to the lack of an explicit expression for the objective function in the constrained optimization problem (i.e., the objective function itself is the SOCP problem defined in Eq. (10)).

To determine when to stop the iterations for solving the optimization problem of Eq. (12), we adopt two criteria, and the satisfaction of either one is assumed to signal the convergence to a solution. Specifically, the optimization ends once (1) the quantity referred to as "first-order optimality" is lower than a tolerance, opt_{tol}, or (2) the norm of the vector containing the changes of nodal positions during an iteration is lower than a tolerance, Δx_{tol}. First-order optimality, described in greater detail by Nocedal and Wright (2006), is a well-known and widely used measure of how close the current solution is to optimal. We use the second criterion to cease iterations when r adaptivity produces only minor perturbations that lead to marginal improvement the computed limit load. The following tolerance values are employed in this work: $\text{opt}_{tol} = 1E^{-2}$ and $\Delta x_{tol} = 1E^{-2}$. The usage of lower tolerances increases the number of iterations but does not noticeably improve the solution. For detailed descriptions of the above stopping criteria and their implementation in MATLAB, the reader is referred to Nocedal and Wright (2006) and The MathWorks, Inc (2018).

5. Element subdivision ($\boldsymbol{h}$ adaptivity)

Once r adaptivity is applied to optimize the limit load and velocity field for a particular mesh topology (element number and connectivity), further improvement of the solution requires either uniformly or selectively refining the mesh. This section proposes a strategy to refine the mesh by selectively dividing elements, such that refinement will only be performed as needed and at locations that potentially improve the solution.

Any subdivision strategy must decide where to refine the mesh based on certain *a posteriori* indicators (i.e., information derived from the current computation). For a rigid block system, a simple indicator is the magnitude of the velocity jump, which is proportional to the integral of strain rate over the infinitesimally thin layers between adjacent elements (Chen, 1975) represented as velocity discontinuities. The magnitude of the velocity jump therefore identifies regions characterized by high strain rate, and mesh refinement in these regions typically has the highest potential for improving the solution. This concept is similar to the adaptive mesh refinement proposed by Martin (2011) for FELA, which attempts to evenly distribute the integral of the maximum shear strain rate over all elements, such that the concentration of elements reflects the intensity of the shearing rate (change of velocity). The specific subdivision criterion postulated in this work is to subdivide elements sharing an edge for which the magnitude of the velocity jump is greater than a tolerance Δv_{tol}, i.e., $\sqrt{\Delta v_i \Delta v_i} \geq \Delta v_{tol}$.

The flow chart within the dashed box of Fig. 2 presents the basic algorithm iterated over all elements to perform the subdivision. This algorithm first filters out elements with nearly zero velocity or small volume through prescribed tolerances v_{tol} and V_{tol}, respectively. The former filtering prevents unnecessary refinement in stationary regions, and the latter contributes to forming the best overall shape of the mechanism, excluding the partition of small elements that tend only to result in small and localized improvement

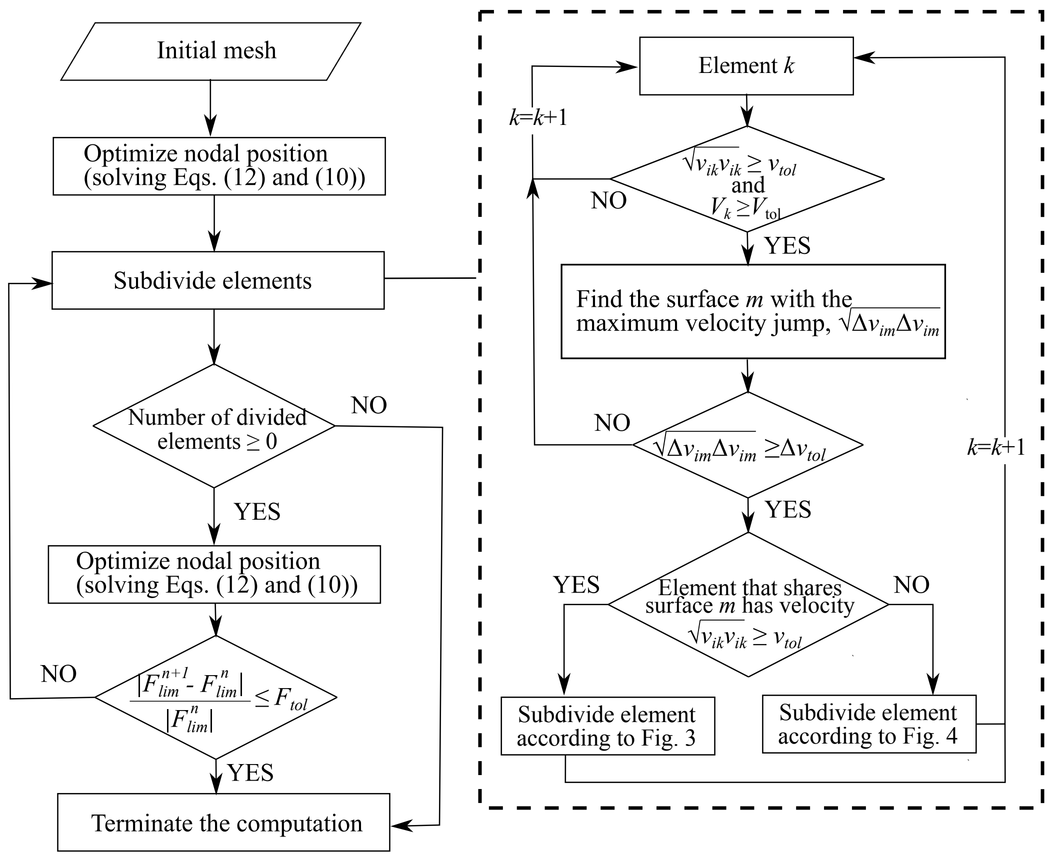


Figure 2: Computation flow chart of the proposed r-h adaptive approach.

of the collapse mechanism. Each element that passes this first screening and has edges with $\sqrt{\Delta v_i \Delta v_i} \geq \Delta v_{tol}$ will be subdivided according to either Fig. 3 or Fig. 4, depending on whether this velocity jump is between two moving elements (i.e., both have velocity greater than v_{tol}), or between a moving element and a stationary region.

As depicted in Fig. 3, when the targeted velocity jump is between two moving elements, we propose two different approaches to subdivide the element corresponding to the subfigures (a) and (b). The adoption of one

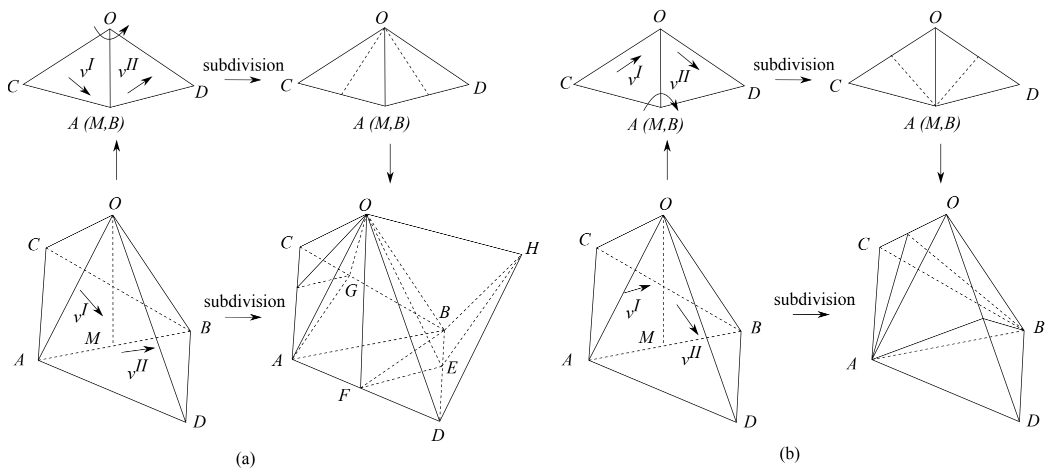


Figure 3: Schematic showing the subdivision of a pair of blocks based on the flow direction of the local velocity field.

of these two alternatives depends on *the flow direction of the local velocity field*. Fig. 3 shows the two possibilities: a flow tending to "rotate" about the point O, as shown in subfigure (a), or rotate about the axis AB, as shown in subfigure (b). Identifying this flow direction is important because regional velocity jumps can be reduced (smoothed) when more elements are added aligning with this direction.

For the scenario shown in Fig. 3(a), the blocks $OABC$ and $OADB$ are divided so that the newly added discontinuities radiate from the point O and bisect the edges AC, BC, AD and BD. Note that for illustration purposes, we have assumed the surface OAB possesses the maximum velocity jump for both blocks; otherwise, only one block is subdivided. Fig. 3(a) also shows that the subdivision of the targeted element $OADB$ adds a new node E to the edge BD, which is shared by an adjacent element $ODBH$. These neighboring elements will automatically be partitioned by new discontinuities passing through the new nodes (e.g., the new discontinuity OEH is created to pass through the node E in Fig. 3(a)). Without such partition of neighboring elements, subsequently changing the positions of the new nodes (e.g., the node E) can lead to interpenetration or gaps between the newly formed elements (e.g., the blocks $OFEB$ and $OFDE$) and those that already existed (e.g., the block $ODBH$). Moreover, subdividing these adjacent blocks ensures that the newly formed discontinuities are connected (e.g., the discontinuities OFE and OEH), thus enabling immediate benefits from the r adaptivity. Due to the fact that only tetrahedral elements are considered, some secondary discontinuities (e.g., the discontinuities OAG and OFB in Fig. 3(a)) are added during the subdivision process. Extending the proposed approach to other element shapes would eliminate this requirement.

When the local velocity field features the characteristics shown in Fig. 3(b), the newly added discontinuities radiate from the axis AB and bisect the edges OC and OD, and there are many possible ways to distinguish the above two different flow directions. The one employed in this work is given by

$$\begin{cases} \text{rotate about} \quad AB \quad \text{if} \quad \Delta v_i r_i < 0 \\ \text{rotate about} \quad O \quad \text{if} \quad \Delta v_i r_i \geq 0 \end{cases} \tag{13}$$

where $\Delta v_i = v_i^I - v_i^{II}$, with v_i^I and v_i^{II} denoting the element velocities pointing toward and away from the shared surface OAB shown in Fig. 3, respectively. The variable r_i in Eq. (13) represents a unit vector pointing from O to M. It is used as a reference direction for distinguishing the direction of the velocity jump. When a pair of elements have velocities that both point toward or away from the interface (OAB in Fig. 3), they will not be divided in the current iteration, due to the ambiguity of the flow direction.

Figure 3 shows only one of three possible permutations, namely the flow direction of the local velocity field can also rotate about the other two pairs: (1) the point A paired with the axis BO and (2) the point B paired with the axis AO. These three possibilities are distinguished by projecting the velocity jump to the three edges of the triangle OAB. The edge with the least projection is the one to which the velocity jump has the greatest perpendicular component, and thereby the one about which the local velocity flow tends to rotate. Mathematically, this criterion can be expressed as

$$\begin{cases} \text{rotate about} \quad O/AB \quad \text{if} \quad |\Delta v_i o_i| \leq \min(|\Delta v_i p_i|, |\Delta v_i q_i|) \\ \text{rotate about} \quad A/BO \quad \text{if} \quad |\Delta v_i p_i| \leq \min(|\Delta v_i o_i|, |\Delta v_i q_i|) \\ \text{rotate about} \quad B/AO \quad \text{if} \quad |\Delta v_i q_i| \leq \min(|\Delta v_i o_i|, |\Delta v_i p_i|) \end{cases} \tag{14}$$

where o_i, p_i and q_i denote vectors along edges AB, BO and AO, respectively.

Elements adjacent to stationary regions are subdivided as illustrated in Fig. 4. Specifically, the element is divided by creating a new discontinuity that radiates from the point O and bisects the edges AC and BC. The

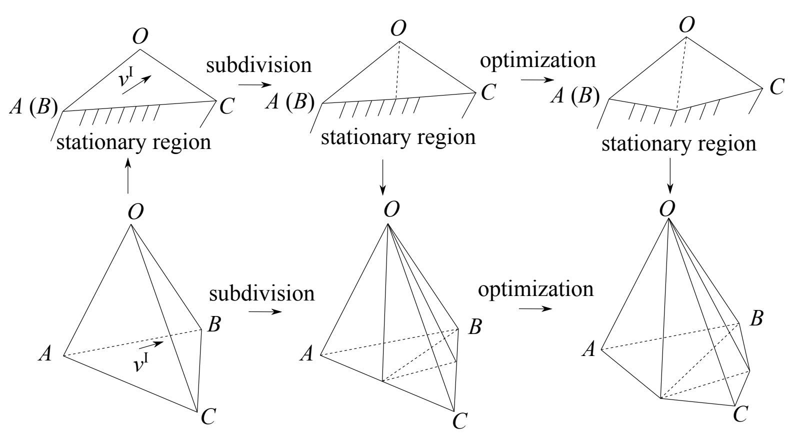


Figure 4: Schematic showing subdivision of a moving block adjacent to stationary region.

decision as to which edges to bisect are determined in a manner similar to Eq. (14). In the rightmost figures, we show the new nodes in their optimized positions (off of plane ABC) to illustrate that this type of subdivision enables an accurate resolution of the boundary between moving material and the stationary region, which is typically a discontinuity whose shape is not known beforehand.

6. Algorithm summary

The complete algorithm for the proposed r-h adaptive method is summarized in the main flow chart of Fig. 2. The computations start by optimizing the nodal positions of the initial mesh. Then, the algorithm repeats the cycle of subdividing elements and adjusting nodal positions, until satisfying either of the following two criteria: (1) the relative improvement of the limit load between two consecutive subdivisions is less than a prescribed tolerance, denoted by F_{tol}, or (2) no element needs to be subdivided. It should be noted that, in the above-mentioned cycle, any h adaptivity step is immediately followed by an r adaptivity step. The reason why we do not allow consecutive h adaptivity steps will be elaborated by the numerical examples detailed in Section 7.

As in any numerical approach, the question arises as to how to select the various tolerances introduced above. For the numerical examples discussed later, trial and error revealed that the following choices of tolerances give satisfactory performance: $F_{tol} = 0.1$, $\Delta v_{tol} = v_0$, $v_{tol} = 0.01v_0$, where v_0 denotes the magnitude of the velocity along the boundary where the limit load is evaluated. Because the volume filtering mechanism described above can potentially stop subdivision prematurely, a small value of $1E^{-3}b^3$ is assigned to the tolerance V_{tol}, where b is the largest dimension of the loading area.

7. Example problems

To explore the performance of the proposed method, three examples are studied: (1) bearing capacity of a square foundation on cohesionless soil or purely cohesive soil; (2) passive uplift resistance of a square, horizontal anchor embedded in cohesionless soil; and (3) passive resistance of a rectangular retaining wall in cohesionless soil.

7.1. Bearing capacity of a square foundation

The limit load for a square surface foundation of width b on cohesionless soil can be expressed as

$$F = \frac{1}{2}\gamma b^3 N_{\gamma s} \tag{15}$$

In Eq. (15), the dimensionless quantity $N_{\gamma s}$ is a function of the internal friction angle ϕ and the interfacial roughness between the footing and the soil. The subscript "s" is used to distinguish this factor, for a square foundation, from the 2D (plane strain) bearing capacity factor commonly denoted as N_γ. Exact values for N_γ were obtained numerically by Martin (2005), who performed detailed calculations based on the method of characteristics and utilized, notably, adaptive subdivision in his approach. In 3D, exact theoretical solutions remain elusive, and $N_{\gamma s}$ in particular is an unknown function. However, upper bounds obtained through limit analysis have been evaluated semi-analytically and numerically (Michalowski, 2001; Krabbenhøft et al., 2008; Lyamin et al., 2007). This work models cohesionless soils by assigning zero-valued cohesion c to the dissipated power (Eq. 10). The unit weight of the soil γ and the footing width b are each assumed to be 1 for ease in interpreting $N_{\gamma s}$. The relative slip between the footing and the soil is prevented (i.e., perfectly rough) in the simulation.

To initiate the computation, one has to guess an initial mesh. For reference, we consider the mechanism constructed by Michalowski (2001) rendered in Fig. 5(a). This mechanism is characterized by a single pyramidal block that moves downward vertically with the foundation and four adjacent regions composed of rigid blocks truncated by conical surfaces. For clarity, Fig. 5 shows only one of the four regions. By comparison, the starting guess considered in this work is extremely simple. It is depicted in Fig. 5(b). Taking advantage of the four-fold symmetry (i.e., OMN shown in Fig. 5(a) represents a 45° slice of the footing), the mesh consists of only three elements (blocks), one directly beneath the foundation and two that are adjacent. Initially, one

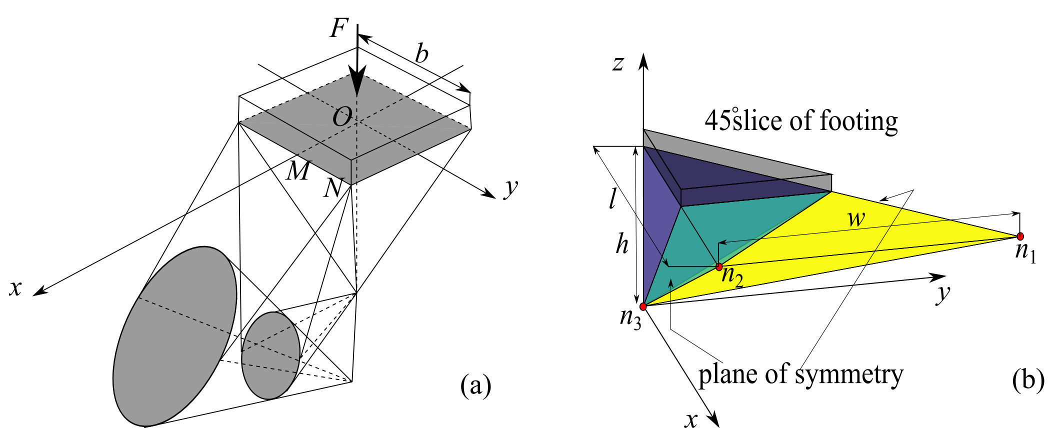


Figure 5: Bearing capacity of a rough rigid square foundation on cohesionless soils: (a) multi-block mechanism (adapted from Michalowski (2001)); (b) initial mesh assumed in the r-h adaptive approach.

has to guess the positions of the nodes n_1, n_2, and n_3 (or equivalently the values of three geometric variables h, l and w in Fig. 5(b)). Throughout the r-h adaptive optimization procedure (Fig. 2), these nodes are constrained to move parallel to the plane of symmetry in which they reside, $y = 0$ or $x = y$, as are any nodes within these planes added through adaptive subdivision. Additionally, the components of velocity normal to the planes are constrained to be zero.

When the friction angle is high, the jump condition given by Eq. (2) becomes increasingly restrictive with respect to finding a kinematically admissible velocity field for a particular mesh. Consequently, the existence of a feasible solution for SOCP becomes sensitive to the mesh geometry, and selecting initial values for the above geometric variables becomes challenging. This issue was resolved by sequentially optimizing the nodal positions while gradually increasing the friction angle. In other words, one can start the computation by (1) introducing a low, fictitious friction angle denoted by ϕ_0, (2) optimizing the nodal positions, and (3) using the optimized mesh as a starting guess to obtain a feasible initial solution for a higher friction angle. The procedure is repeated until the true friction angle is reached. In this work, the starting guess in all cases was $h = b/2$, $w = b$, and $l = b$ with $\phi_0 = 10°$.

The solid line in Fig. 6 shows the computed values of $N_{\gamma s}$ for $\phi = 35°$ as they vary for each iteration of the SQP algorithm utilized within the proposed r-h adaptive solution procedure to solve Eq. (12). The figure shows that the computed upper bound on $N_{\gamma s}$ rapidly decreases as the iteration number increases, highlighting the sensitivity of the solution to the mesh,

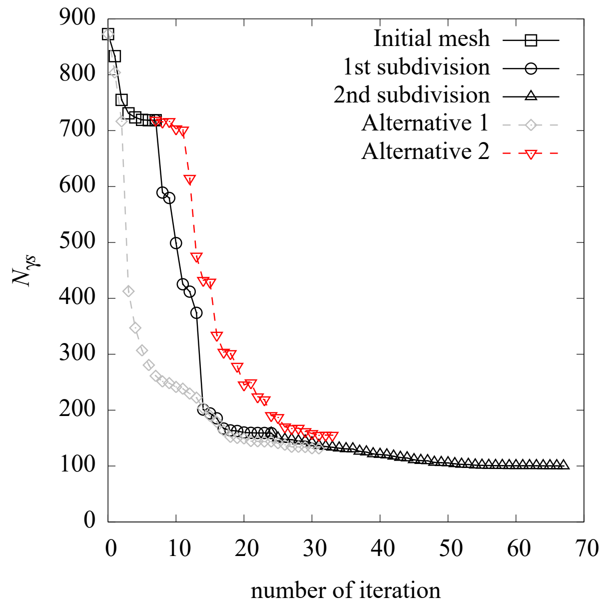
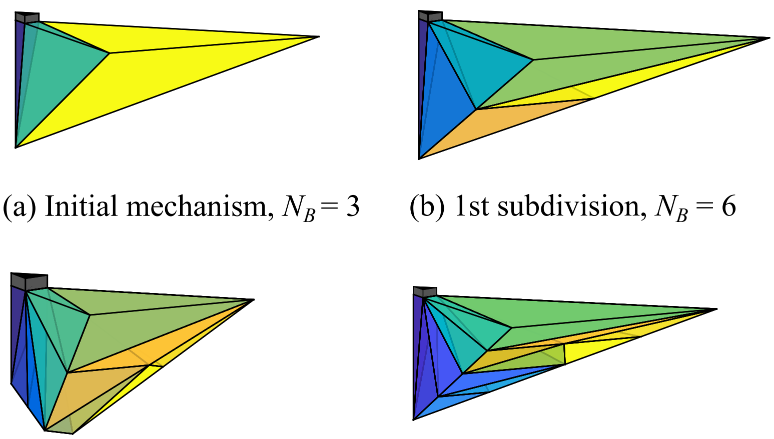


Figure 6: Variation of the $N_{\gamma s}$ value as a function of the iteration numbers in the non-linear optimization ($\phi = 35°$).

and thus also revealing the effectiveness of r adaptivity. In this example, the method resulted in two subdivisions. The initial mesh and the meshes corresponding to these subdivisions are presented in Figs. 7(a)-(c), wherein the number of elements after each subdivision is also provided. Prior to each subdivision, the convergence curve becomes flat, signaling that better upper bounds cannot be reached for the current mesh. Through the use of h adaptivity, the computed limit load can be further reduced, and a faster convergence rate can be recovered (e.g., 1st subdivision in Fig. 6). The reason why h adaptivity is effective is revealed in Fig. 7. Comparing the initial mesh to those obtained after subdivision, the approach enables the creation of more velocity discontinuities radiating outward from the edge



(a) Initial mechanism, $N_B = 3$ (b) 1st subdivision, $N_B = 6$

(c) 2nd subdivision, $N_B = 14$ (d) Alternative 1, $N_B = 14$

Figure 7: Mesh before and after element subdivisions ($\phi = 35°$).

of the footing. Moreover, the lowermost part of the collapse mechanism is gradually divided in a way that the above radial discontinuities can extend all the way to the boundary of the region of failing (moving) soil. Both features are important in forming a radial shearing zone, which accommodates the rotation of the principal directions of strain.

Figure 6 highlights the fact that the rate of improvement in the solution generally diminishes as the r-h adaptive iterations proceed. This response may be attributed to two possible explanations. The first hypothesis is that the smaller element sizes obtained through h adaptivity constrain the magnitude of nodal position changes that can occur in the optimization, due to the imposed non-linear constraints requiring no interpenetration between elements (see Eq. (12)). This reduces the amount that nodes are able to perturb around their current positions, thereby demanding more r-adaptive iterations to achieve a better mechanism. The second hypothesis is that, as the current mechanism is closer to the optimum, the optimization algorithm adopts smaller step size (i.e., nodal perturbation during one iteration), and consequently the rate of improvement is reduced.

To test these two hypotheses, we consider an alternative initial mesh with the same overall geometry as the original one (Fig. 7(a)) but with the same connectivity and number of elements as in the final solution (Fig. 7(c)). Compared with the original starting guess, this new initial mechanism (Fig. 7(d)) simply has smaller initial element sizes. The dashed line in Fig. 6 designated by "Alternative 1" shows that the r adaptive iterations are more effective for the new initial mesh with smaller element sizes. This reveals that the deterioration in the effectiveness of r adaptivity is not related to the number and size of elements but rather due to the fact that an optimal mechanism is approached (the second of the two hypotheses above). One might infer from the above discussion that starting from a more refined mesh is generally more effective, given that better results are achieved with fewer iterations. However, this is not the case, since the refined solution with element edges (velocity discontinuities) placed at strategic locations is known only after refinements are obtained through iterations of r-h adaptivity.

We use the data corresponding to "Alternative 2" in Fig. 6 to illustrate why an r adaptive step is employed immediately following any h adaptive step, as described in Section 6. In Alternative 2, two consecutive h adaptive steps are performed on the initial mesh depicted in Fig. 7(a). The nodal positions of this refined mesh are then optimized using r adaptivity. Figure 6 shows that after multiple h adaptive steps, r adaptivity becomes less efficient compared with the proposed algorithm. This can be explained as follows. When only subdividing elements without optimizing nodal positions, the elements on both sides of the new discontinuities have the same velocity as if the original elements have yet to be subdivided, and there are no velocity jumps across these new discontinuities. These new discontinuities with zero velocity jumps do not provide effective information regarding how to refine the mesh (see Section 5). This analysis shows that simply increasing the number of elements often does not lead to an improved solution. It underscores the merit of the proposed approach, which starts from a simple mesh that is progressively refined through combined r-h adaptivity.

Figure 8 presents the $N_{\gamma s}$ values computed with the proposed method for friction angles from $\phi = 15°$ to $35°$. For the purpose of comparison, the existing semi-analytical solution of Michalowski (2001) and the FELA results of

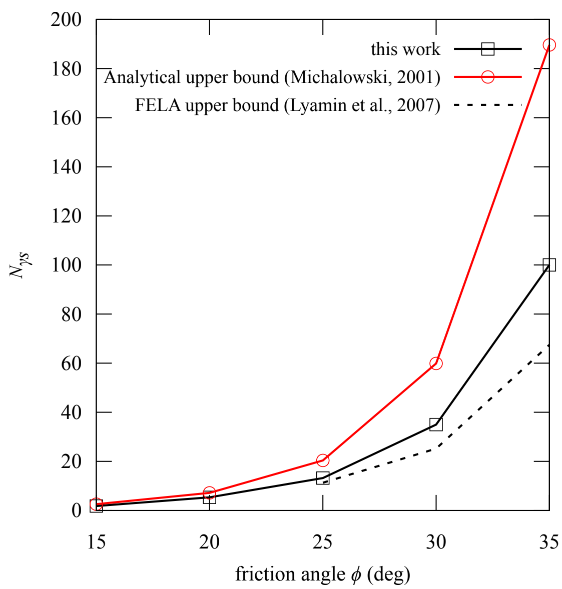


Figure 8: Comparison of $N_{\gamma s}$ values computed from the proposed method and existing solutions.

Lyamin *et al.* (2007) are included in the figure. Figure 8 shows that the proposed method gives a better (smaller) solution than the analytical approach, with the improvement increasing as the friction angle grows. On the other hand, the $N_{\gamma s}$ values computed in the present study are larger than the ones given by FELA, with the difference again tending to increase as the friction angle increases. Such a discrepancy between these two methods might be attributed the the continuous deformation allowed within elements in FELA. Through the use of rigid elements, the implementation presented in this work is potentially restrictive in the manner in which it accommodates the dilation of soils with large friction angles. Nevertheless, the ability of such a simple approach, and relatively simple collapse mechanism, to capture reasonable values of the limit load for such a challenging problem is remarkable.

Figure 9 compares the computed failure mechanisms for different friction

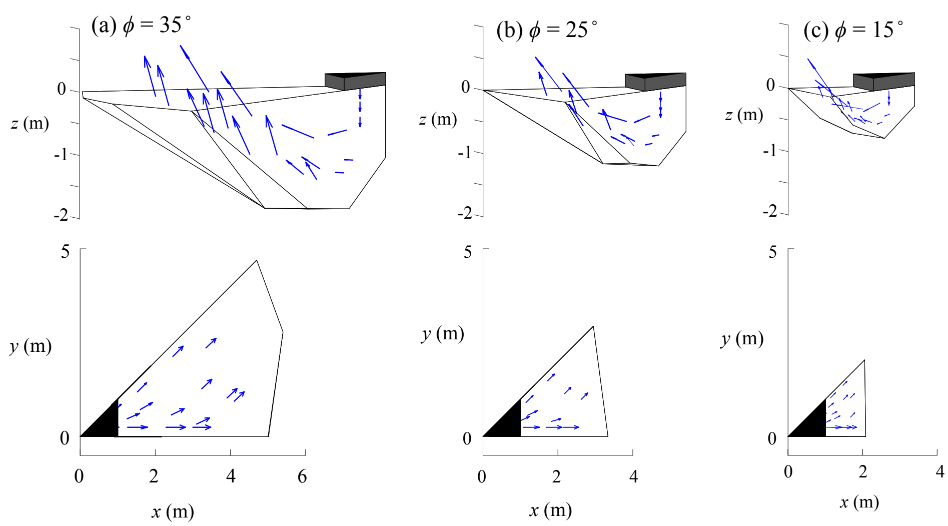


Figure 9: Optimal failure mechanism beneath the square foundation computed with the proposed approach: (a) $\phi = 35°$; (b) $\phi = 25°$; (c) $\phi = 15°$.

angles. As a matter of clarity, some block edges are removed from the figure. It should be emphasized that while all three mechanisms start from the same guess in terms of the initial mesh (Fig. 5(b)), they automatically evolve depending on the friction angle of the material. For larger friction angles, the failure mechanism extends both horizontally and vertically a larger distance compared to solutions with lower friction angles.

To further test the proposed method, especially against existing techniques, the problem of a square foundation on cohesive soil is analyzed. The limit load for this problem, first considered by Shield and Drucker (1953), can be expressed in terms of the soil cohesion c as

$$F = cb^2 N_{cs} \tag{16}$$

where N_{cs} is a constant. Computations were completed in the same manner as for $N_{\gamma s}$, using the same initial guess for the mesh as described above (Fig. 5(b); $h = b/2$, $w = b$, and $l = b$). The unit weight of soils is assumed to be zero and the cohesion c is equal to 1.

Table 1 compares the N_{cs} values computed in this study to those obtained semi-analytically (Michalowski, 2001), using FELA (Vicente da Silva and Antão, 2008), and using DLO (Hawksbee et al., 2013). The recorded or reported computation time for each numerical method is also included. Calculations in this study were completed on a PC equipped with an Intel i7-4790 processor (3.6 GHz; 4 cores) and 8 GB memory. Results from FELA were obtained by distributing computations over 5 or 18 PCs, where each PC was equipped with a single core processor clocked at 3.0 GHz (Intel Pentium IV) and 512 MB memory. The DLO computations were performed on a workstation equipped with an AMD Opteron 6140 processor (2.6 GHz; 8

Table 1: Comparison of computed N_{cs} values by different methods with corresponding computation costs (wall-clock time: the total processing time, including time spent on pre-processing, kernel computation through the FMINCON function and MOSEK, and post-processing; MOSEK time: the processing time spent on solving the second-order cone programming through MOSEK).

Analytical upper bound	FELA upper bound		DLO			This work			
N_{cs}	N_{cs}	wall-clock times (s)	node spacing	N_{cs}	MOSEK times (s)	subdivision	N_{cs}	wall-clock times (s)	MOSEK times (s)
6.56	6.05	2000	1/2	6.52	0.02	0	8.27	0.96	0.2
		to	1/4	6.41	13	1	6.67	2.0	0.4
		15000	1/6	6.22	6400	2	6.44	4.4	0.9

cores) and 8 GB memory, and only the CPU times for SOCP with Mosek were reported. Because the basic computation unit in the proposed r-h adaptive approach is solving Eq. (10) using SOCP, the CPU times for executing Mosek are separated from the total wall-clock times. Due to differences in the hardware and the particulars of programming (e.g., language and code optimization), the computation times reported in Table 1 are merely indicators of the computation cost rather than strictly comparable performance measures.

Table 1 shows that DLO and the r-h adaptive approach provide reasonably accurate estimates of the limit load (better than the analytical solution), and that FELA gives the least upper bound (best estimate of the limit load).

Compared to DLO and FELA, for which accuracy and computation time depend strongly on the element size or grid spacing, the r-h adaptive approach displays a significant improvement in the computed limit load without an exorbitant increase in computational cost. This difference can be attributed to the fact that uniform mesh or grid refinement tends to add a large number of additional unknowns that do not contribute towards improving the solution. Finally, we note that the CPU times for running Mosek in this work are only a small portion of the total times, thus suggesting that the reported computational times can be potentially reduced by utilizing more efficient optimization schemes and programming languages for the non-linear optimization problem posed by r adaptivity (Section 4).

7.2. Uplift resistance of a plate anchor in cohesionless soil

With reference to the collapse mechanism considered by Murray and Geddes (1987), Fig. 10(a) illustrates the problem of a horizontal anchor problem embedded at depth h. The anchor is square with sides of length, b, and the material is assumed to be cohesionless. The status of "immediate breakaway" is considered, which implies that the underside of the anchor loses contact with the soil. The ultimate uplift force F is expressed as

$$F = \gamma h b^2 N_{\gamma b} \tag{17}$$

The factor $N_{\gamma b}$ is referred to as the anchor break-out factor, and its value depends on ϕ, the ratio of the embedment depth to the anchor width (h/b), and friction at the soil-anchor interface. For a fixed friction angle of $\phi = 30^\circ$, this example considers the $N_{\gamma b}$ factors corresponding to varying values of h/b. Here, as in the previous example, we assume zero-valued cohesion to

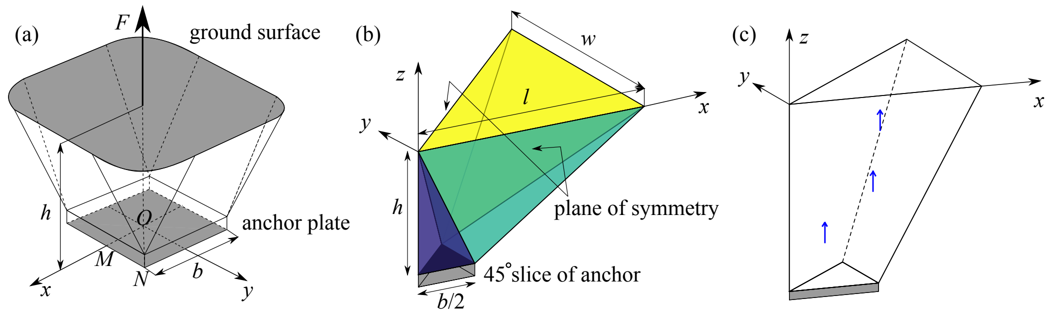


Figure 10: Uplift of an anchor in cohesionless soil: (a) collapse mechanism considered by Murray and Geddes (1987); (b) initial mesh used in the r-h adaptive approach; (c) typical collapse mechanism computed with the r-h adaptive approach.

model cohesionless soils. The unit weight of the soil γ and the anchor width b are each assumed to be 1. Above the anchor, a perfectly rough interface is simulated by eliminating relative movement between the anchor and the soil.

Fig. 10(b) depicts the initial mesh selected for the r-h adaptive approach. As in the previous example, symmetry is invoked to reduce the model to a 45° slice of the anchor (i.e., slice OMN in Fig. 10(a)), where the planes $y = 0$ and $x = y$ represent the planes of symmetry. The initial mesh is again one of the simplest conceivable, and it consists of three elements. The geometric variables l and w are initially assumed to be $2h$, which leads to feasible initial solutions for all embedment ratios.

Figure 11 compares the $N_{\gamma b}$ values computed with the r-h adaptive approach to values obtained in previous works: those obtained with the analytical solution of Murray and Geddes (1987) and the 3D DLO analysis of Hawksbee et al. (2013). Satisfactory agreement between the three methods can be observed for both shallow and deep embedment. Figure 10(c) presents

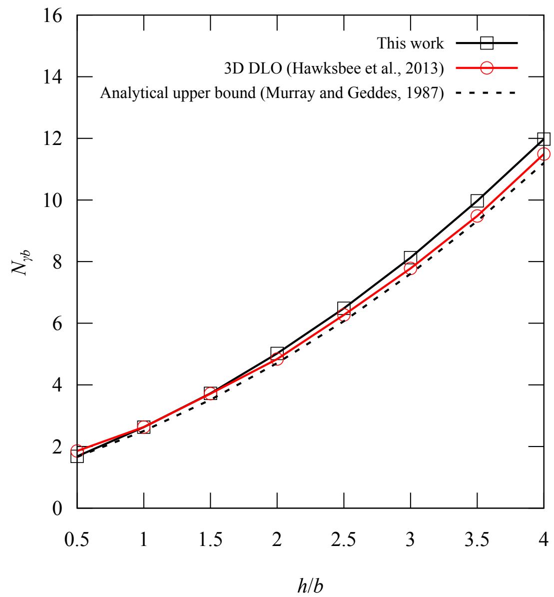


Figure 11: $N_{\gamma b}$ values computed with the r-h adaptive approach compared to existing solutions.

a typical collapse mechanism computed by this work. The mechanism is characterized by single active velocity discontinuity that extends from the edge of the anchor to the ground surface and bounds a plug of material that moves upward with the anchor. This mechanism is similar to the one constructed by Murray and Geddes (1987) (Fig. 10(a)), but it differs with respect to the the conical surfaces assumed at the edges of the collapse mechanism. With these cone-shaped edges, the soil volume lifted by the anchor is reduced, and consequently a slightly lower (better) upper-bound solution is obtained, as shown in Fig. 11.

One possible cause for the above mismatch is that no subdivision step is performed for this anchor problem, as all velocity jumps are below the tolerance for triggering h adaptivity (i.e., $\Delta v_{tol} \leq v_0$, where v_0 in this example is the anchor velocity). Accordingly, we lowered the tolerance Δv_{tol} to $0.1v_0$ to explore whether the Murray and Geddes (1987)'s solution can be recovered by refining the mesh. Figure 12 compares the optimized collapse mechanism based on the initial mesh and the refined one. It can be seen that because the initial mesh produces a uniform velocity field across elements (i.e., velocity jumps between elements are zero), only the blocks at the boundary of the failing soil volume are subdivided. In other words, this subdivision is based exclusively on the velocity jumps between moving blocks (i.e, the blocks $OBDA$ and $OCDB$ in Fig. 12(a)) and the assumed stationary region. Whereas a discontinuity passing the nodes D and O and insects the edge AB could lead to an improved mechanism, Figure 12(b) shows that the proposed subdivision strategy adds new discontinuities that intersect with the failure plane $ABCD$ and do not help in forming a more critical mechanism. As a consequence, the elements and discontinuities introduced in the refined mesh

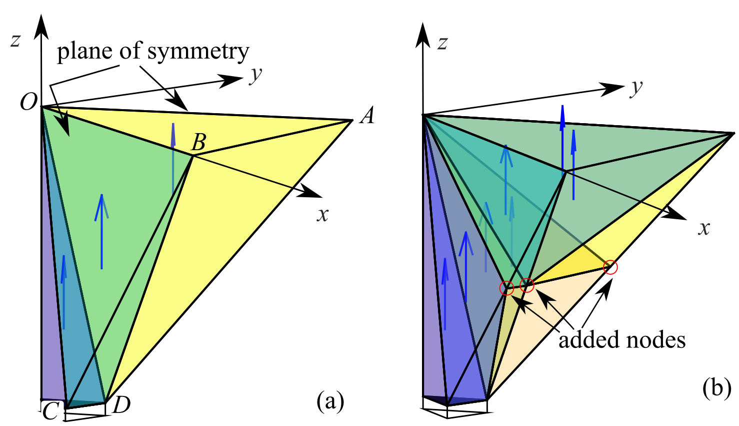


Figure 12: Collapse mechanism after optimizing nodal positions: (a) initial mesh; (b) after first subdivision.

do not alter the optimized collapse mechanism or the computed limit load.

7.3. Rectangular wall in cohesionless soil

As a final example, the r-h adaptive approach is applied to compute the limit load on a rectangular retaining wall in cohesionless soil. The problem is illustrated in Fig. 13(a), which also depicts the collapse mechanism assumed by Soubra and Regenass (2000). The width and height of the wall are denoted by b and h, respectively, and the wall is assumed to move laterally into the soil (passive condition). The passive force on the wall at collapse can be expressed as

$$F = \frac{1}{2}\gamma b h^2 K_{\gamma p} \tag{18}$$

where $K_{\gamma p}$ is the so-called passive earth pressure coefficient. Generally, the passive resistance depends on the mode of wall movement, and in particular whether it translates, rotates, or moves with combined translation and rotation (Widuliński et al., 2011). Here, only translational movement is considered.

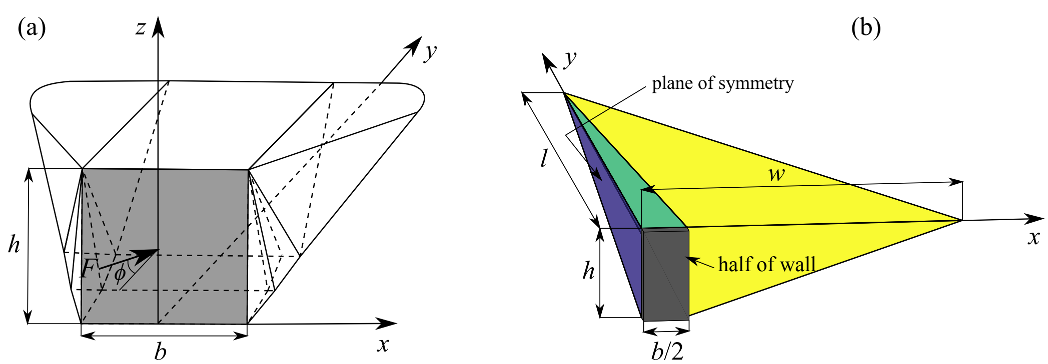


Figure 13: Collapse mechanism for passive failure of a rectangular retaining wall in cohesionless soil: (a) truncated multi-block mechanism (adapted from Soubra and Regenass (2000)); (b) initial mesh assumed in the r-h adaptive approach.

This example models cohesionless soils in the same manner as those discussed earlier. The unit weight γ is assumed to be 1, and both the wall width b and height h are assumed to be 2. Unlike the previous two examples, the perfectly rough interface between soils and the retaining wall is modeled as a velocity discontinuity, whose jump condition is characterized by friction angle ϕ. This change is made in accordance with the assumption in the work of Soubra and Regenass (2000), thus enabling a direct comparison.

Figure 13(b) shows the starting mesh used to initiate the computation. Considering that $x = 0$ is a plane of symmetry, only half of the wall is modeled. The geometric variables l and w are initially assumed to be $2h$, and are adjusted to $l = 3.53h$ and $w = 7.20h$, for all values of ϕ, by the sequential optimization discussed in Section 7.1. In this case the direction of the force F is not horizontal but inclined at an angle ϕ with respect to the direction normal to the wall (see Fig. 13(a)).

Figure 14 compares the $K_{\gamma p}$ values computed in this study and to those assessed using the analytical solution proposed by Soubra and Regenass (2000). The analytical solution corresponds to the failure mechanism shown in Fig. 13(a), which consists of multiple blocks truncated by portions of circular cones. The methods provide very close results for small friction angles. As the friction angle increases, the r-h adaptive approach gives lower (better) estimates of the limit load.

The collapse mechanisms assessed through the r-h adaptive approach for large and small friction angles are presented in Fig. 15. While both cases start from the same initial mechanism, the proposed adaptive approach allows for the mechanism to extend to greater depth and horizontal distance as the

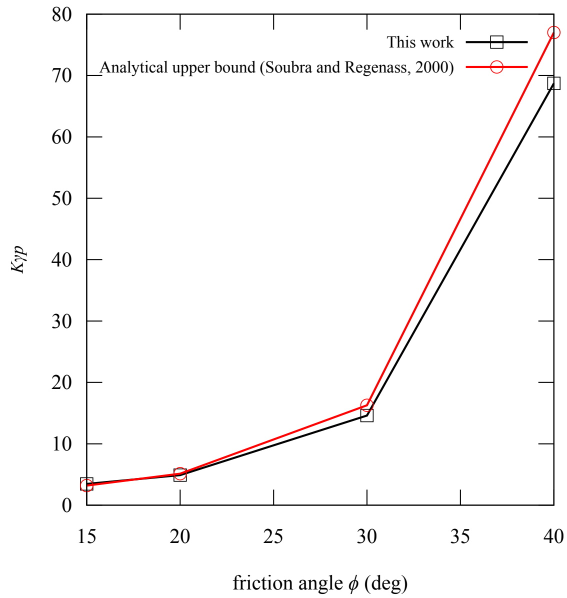


Figure 14: Comparison of the passive earth pressure coefficients $K_{\gamma p}$ computed by this work and the analytical solution of Soubra and Regenass (2000).

friction angle grows, as in the solution of Soubra and Regenass (2000).

8. Discussion

Table 2 summarizes the computational cost of the r-h adaptive approach for the three examples considered. The table includes the number of rigid blocks (N_B), the number of nodal position components subjected to optimization (DOF), wall-clock times, and the CPU times required to run MOSEK. Such information is organized for both the initial mesh configuration and those after h adaptivity steps. For these examples, the r-h adaptive approach displays promising computational efficiency. The maximum computation time is no more than 30 seconds, observed in the square footing

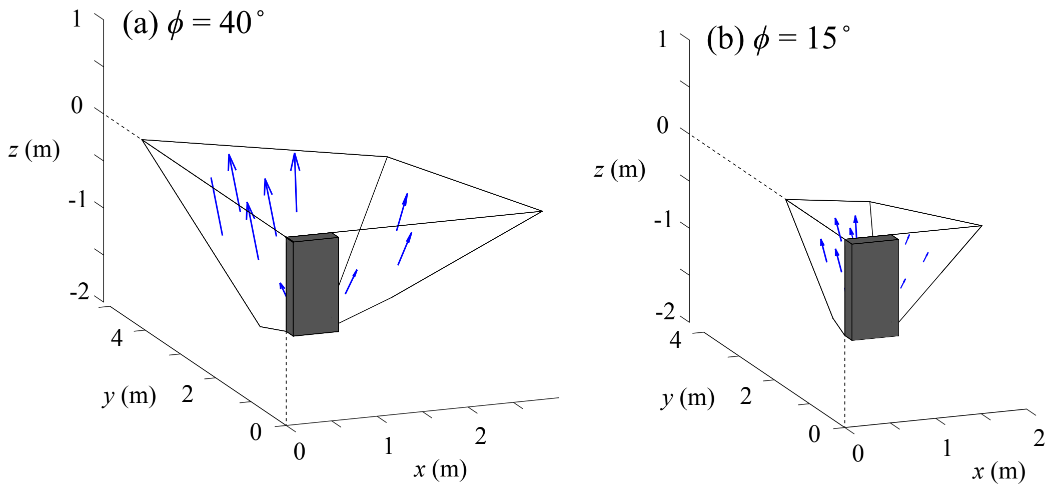


Figure 15: Collapse mechanisms computed with the r-h adaptive approach: (a) $\phi = 40°$; (b) $\phi = 15°$.

Table 2: Computational cost for the three examples.

Example	subdivision	N_B	N_{DOF}	wall-clock time (s)	MOSEK time (s)
square footing	0	3	3	0.8 to 1.2	0.1 to 0.3
	1	5 to 6	5 to 7	2.0 to 4.3	0.4 to 0.6
	2	11 to 14	11 to 18	4.4 to 21.2	0.9 to 4.2
retaining wall	0	3	2	1.7 to 2.1	0.6 to 0.7
	1	10 to 11	10 to 14	1.8 to 9.4	0.4 to 1.6
square anchor	0	3	2	1.8 to 2.2	0.7 to 0.8

case. The fact that the maximum MOSEK running time is around 4 seconds suggests that the approach could also be accelerated by formulating a more efficient strategy to solve the non-linear optimization problem of Eq. (12), rather than using the FMINCON solver available in MATLAB.

While the above computation times are promising, they are not yet sufficiently small to enable highly efficient sequential kinematic analysis for 3D applications, one of the underlying objectives of this work. The computational demands of the proposed approach can be traced to the fact that a forward numerical differentiation is employed to compute the gradient of the objective function in the non-linear optimization. In other words, the objective function (the SOCP problem of Eq. 10) is called $DOF + 1$ times to obtain the gradient. Therefore, computing the gradient consumes a significant amount of time when DOF, corresponding to the number of nodal positions, becomes large. To improve the computational efficiency, a future refinement of the current work could be to approximate rather than directly compute the gradient. For instance, the objective function can be linearized with respect to its unknowns (cf. Hambleton and Sloan, 2013; Milani and Lourenço, 2009), thus rendering an approximated but analytical form of the gradient.

Compared with previous works on r adaptivity, an important contribution of this work is the adaptive subdivision, which automatically changes the topological connectivity of blocks based on velocity jumps between blocks. Table 3 summarizes the computed limit loads under initial mesh configuration and after subsequent subdivisions. It can be seen that the calculated upper bounds significantly decrease as the mesh is gradually refined, thus suggesting that element subdivision based on the velocity jump is effective in improving the collapse mechanism. Table 3 reports the computed results only up to 2 subdivisions. The reason is that further mesh refinements only lead to marginal improvement on the collapse loads. For example, additional mesh refinement only decreases the computed $N_{\gamma s}$ value in the square foot-

Table 3: Limit loads computed by the initial mesh and after element subdivision (all reported values are obtained after optimizing nodal positions).

Numerical example	ϕ (°)	initial mesh	1st subdivision	2nd subdivision
square footing $N_{\gamma s}$ or N_{cs}	35	718.6	159	100.2
	30	150.5	50.9	35.0
	25	41.9	17.9	13.2
	20	13.6	6.6	5.4
	15	4.8	2.4	1.8
	0	8.3	6.7	6.4
retaing wall $K_{\gamma p}$	40	699.7	68.7	-
	30	23.9	14.6	-
	20	5.8	4.9	-
	15	3.5	3.5	-

ing example by less than 5%, while the coefficient $K_{\gamma p}$ in the retaining wall problem remains unchanged even more subdivisions are performed.

The fact r-h adaptive approach eventually reaches a limit of no improvement can be attributed to two reasons. First, as demonstrated explicitly in Section 7.2, the proposed subdivision scheme does not always lead to an improvement in the solution. Indeed, the development of a more sophisticated subdivision strategy is an matter for future investigation. Such future algorithms can be devised by (1) identifying other useful indicators that flag the regions to be refined and (2) devising effective methods to subdivide elements so that discontinuities can be added at strategic locations. The second reason is that the algorithm used to solve the non-linear optimization is only a local optimizer, and the solution is susceptible to being trapped at a point that is a local rather than global optimum. The likelihood of this occurring increases as elements are are subdivided, since the the number of unknowns (nodal positions) handled by the optimization is higher. As a part of future work, global optimization techniques (e.g., genetic algorithm) can be employed to resolve this potential limitation.

The initial mesh used as a starting guess in the proposed algorithm also plays a significant role in the accuracy of the computed solution and whether or not a global minimum can be attained. As made evident in the results shown in Table 1, 3D DLO with a coarse grid may provide a reasonable estimate of the limit loads at low cost, thus representing an encouraging approach to systematically define initial meshes that can subsequently be refined using the approach proposed in this work.

9. Conclusions

We propose an r-h adaptive kinematic approach for computing collapse mechanisms and limit loads in 3D problems. Considering a velocity field consisting of rigid elements (blocks) separated by zero-thickness velocity discontinuities, this method progressively improves the collapse mechanism and bound on the limit load by successively adjusting the element nodal positions (r adaptivity) as well as the element number and connectivity (h adaptivity). Examination of the proposed technique through examples shows that when the optimal mechanism is relatively simple, satisfactory limit loads can be obtained solely by optimizing nodal positions (i.e., the locations of velocity discontinuities), even if a simple mesh is assumed. However, when the optimal collapse mechanism becomes more intricate, adding discontinuities at critical locations becomes crucial for the performance of r adaptivity. The subdivision scheme proposed in this work automatically splits existing elements with velocity jumps greater than a specified threshold, adding new elements so that velocity jumps can be further reduced through r adaptivity. This approach allows for the initiation of calculations from a very simple mesh to which new discontinuities are progressively added at critical locations, a paradigm that gives demonstrably high efficiency and may yield higher efficiencies with future refinements.

To further speed up computations and enable efficient sequential kinematic analysis, wherein a full process of deformation is simulated through a series of kinematic limit analysis computations, the proposed method can be improved by pursuing alternatives to solving the non-linear optimization of Eq. (12), devising more effective subdivision schemes, and developing a systematic means of defining the initial mesh. These future refinements represent important steps towards efficiently simulating large deformation problems, especially those involving cohesionless soils, that are extremely challenging to model by any other means.

10. Acknowledgements

Aspects of this work were supported by the Australian Research Council through the award of Discovery Early Career Researcher Award (DE160100328) and by the National Science Foundation (CMMI-1742849). This support is gratefully acknowledged.

References

Antão, A. N., Santana, T. G., Vicente da Silva, M., da Costa Guerra, N. M., 2011. Passive earth-pressure coefficients by upper-bound numerical limit analysis. Canadian Geotechnical Journal 48 (5), 767–780.

Borges, L., Feijó, R., Zouain, N., 1999. A directional error estimator for adaptive limit analysis. Mechanics Research Communications 26 (5), 555–563.

Borges, L., Zouain, N., Costa, C., Feijóo, R., 2001. An adaptive approach to limit analysis. International Journal of Solids and Structures 38 (10), 1707–1720.

Chen, W.-F., 1975. Limit Analysis and Soil Plasticity. Elsevier.

Ciria, H., Peraire, J., Bonet, J., 2008. Mesh adaptive computation of upper and lower bounds in limit analysis. International journal for numerical methods in engineering 75 (8), 899–944.

Davis, E. H., 1968. Theories of plasticity and the failure of soil masses. In: Lee, I. K. (Ed.), Soil Mechanics: Selected Topics. Butterworths, Sydney, pp. 341–380.

Davis, E. H., Booker, J. R., 1971. The bearing capacity of strip footing from the standpoint of plasticity theory. In: Proc. the First Australia-New Zealand Conference on Geomechanics.

Drescher, A., 1991. Analytical Methods in Bin-Load Analysis. Vol. 36 of Developments in Civil Engineering. Elsevier.

Drescher, A., Detournay, E., 1993. Limit load in translational failure mechanisms for associative and non-associative materials. Géotechnique 43 (3), 443–456.

Dunne, H. P., Martin, C. M., 2017. Capacity of rectangular mudmat foundations on clay under combined loading. Géotechnique 67 (2), 168–180.

Godwin, R., O'Dogherty, M., 2007. Integrated soil tillage force prediction models. Journal of Terramechanics 44 (1), 3–14.

Griffiths, D. V., Marquez, R. M., 2007. Three-dimensional slope stability analysis by elasto-plastic finite elements. Géotechnique 57 (6), 537–546.

Hambleton, J. P., 2017. Earthmoving through the lens of geotechnical engineering. In: 6th International Young Geotechnical Engineers Conference (iYGEC6), Seoul, Korea, Sept. 17-22.

Hambleton, J. P., Drescher, A., 2012. Approximate model for blunt objects indenting cohesive-frictional materials. International Journal for Numerical and Analytical Methods in Geomechanics 36 (3), 249–271.

Hambleton, J. P., Sloan, S. W., 2013. A perturbation method for optimization of rigid block mechanisms in the kinematic method of limit analysis. Computers and Geotechnics 48, 260–271.

Hambleton, J. P., Sloan, S. W., 2016. A simplified kinematic method for 3D limit analysis. Applied Mechanics and Materials 846, 342–347.

Hambleton, J. P., Stanier, S. A., White, D. J., Sloan, S. W., 2014. Modelling ploughing and cutting processes in soils. Australian Geomechanics 49 (4), 147–156.

Hawksbee, S., Smith, C., Gilbert, M., 2013. Application of discontinuity layout optimization to three-dimensional plasticity problems. Proc R Soc A 469 (20130009).

He, L., Gilbert, M., 2016. Automatic rationalization of yield-line patterns identified using discontinuity layout optimization. International Journal of Solids and Structures 84, 27–39.

Hettiaratchi, D. R. P., Reece, A. R., 1974. The calculation of passive soil resistance. Géotechnique 24, 289–310.

Johnson, D., 1995. Yield-line analysis by sequential linear programming. International Journal of Solids and Structures 32 (10), 1395–1404.

Kashizadeh, E., Hambleton, J. P., Stanier, S. A., 2014. A numerical approach for modelling the ploughing process in sands. In: Proc. 14th International Conference of the International Association for Computer Methods and Advances in Geomechanics, Kyoto, Japan, Sept. 22-25. pp. 159–164.

Kong, D., Martin, C. M., Byrne, B. W., 2017. Sequential limit analysis of pipe–soil interaction during large-amplitude cyclic lateral displacements. Géotechnique 68 (1), 1–12.

Krabbenhoft, K., Karim, M. R., Lyamin, A. V., Sloan, S. W., 2012. Associated computational plasticity schemes for nonassociated frictional materials. International Journal for Numerical Methods in Engineering 90 (9), 1089–1117.

Krabbenhøft, K., Lyamin, A. V., Sloan, S. W., 2008. Three-dimensional mohr–coulomb limit analysis using semidefinite programming. Communications in Numerical Methods in Engineering 24 (11), 1107–1119.

Lyamin, A. V., Salgado, R., Sloan, S. W., Prezzi, M., 2007. Two-and three-dimensional bearing capacity of footings in sand. Géotechnique 57 (8), 647–662.

Lyamin, A. V., Sloan, S. W., 2002a. Lower bound limit analysis using non-linear programming. International Journal for Numerical Methods in Engineering 55 (5), 573–611.

Lyamin, A. V., Sloan, S. W., 2002b. Upper bound limit analysis using linear finite elements and non-linear programming. International Journal for Numerical and Analytical Methods in Geomechanics 26 (2), 181–216.

Lyamin, A. V., Sloan, S. W., Krabbenhøft, K., Hjiaj, M., 2005. Lower bound limit analysis with adaptive remeshing. International Journal for Numerical Methods in Engineering 63 (14), 1961–1974.

Martin, C. M., 2005. Exact bearing capacity calculations using the method of characteristics. In: Proc. IACMAG. Turin. pp. 441–450.

Martin, C. M., 2011. The use of adaptive finite-element limit analysis to reveal slip-line fields. Géotechnique Letters 1 (4-6), 23–29.

Martin, C. M., Makrodimopoulos, A., 2008. Finite-element limit analysis of mohr–coulomb materials in 3D using semidefinite programming. Journal of Engineering Mechanics 134 (4), 339–347.

Mary, B. C. L., Maillot, B., Leroy, Y. M., 2013. Deterministic chaos in frictional wedges revealed by convergence analysis. International Journal for Numerical and Analytical Methods in Geomechanics 37 (17), 3036–3051.

Michalowski, R., Drescher, A., 2009. Three-dimensional stability of slopes and excavations. Géotechnique 59 (10), 839–850.

Michalowski, R. L., 2001. Upper-bound load estimates on square and rectangular footings. Géotechnique 51 (9), 787–798.

Milani, G., 2015. Upper bound sequential linear programming mesh adaptation scheme for collapse analysis of masonry vaults. Advances in Engineering Software 79, 91–110.

Milani, G., Lourenço, P. B., 2009. A discontinuous quasi-upper bound limit analysis approach with sequential linear programming mesh adaptation. International Journal of Mechanical Sciences 51 (1), 89–104.

Mosek, 2015. The mosek optimization toolbox for MATLAB manual. Version 7.1 (Revision 28), 17.

Muñoz, J. J., Bonet, J., Huerta, A., Peraire, J., 2009. Upper and lower bounds in limit analysis: adaptive meshing strategies and discontinuous loading. International Journal for Numerical Methods in Engineering 77 (4), 471–501.

Muñoz, J. J., Hambleton, J., Sloan, S. W., 2018. R-adaptivity in limit analysis. In: Barrera, O., Cocks, A., Ponter, A. (Eds.), Advances in Direct Methods for Materials and Structures. Springer International Publishing, Cham, pp. 73–84.

Murray, E. J., Geddes, J. D., 1987. Uplift of anchor plates in sand. Journal of Geotechnical Engineering 113 (3), 202–215.

Nocedal, J., Wright, S. J., 2006. Numerical optimization. Springer.

Shield, R. T., Drucker, D. C., 1953. The application of limit analysis to punch-indentation problems. Journal of Applied Mechanics 20, 453–460.

Sloan, S. W., 1995. Limit analysis in geotechnical engineering. In: Haberfield, C. M. (Ed.), Proc. Ian Boyd Donald Symposium on Modern Developments in Geomechanics, Melbourne, Australia, June 7.

Sloan, S. W., 2013. Geotechnical stability analysis. Géotechnique 63 (7), 531.

Smith, C., Gilbert, M., 2007. Application of discontinuity layout optimization to plane plasticity problems. Proc R Soc A 463, 2461–2484.

Soubra, A.-H., Regenass, P., 2000. Three-dimensional passive earth pressures by kinematical approach. Journal of Geotechnical and Geoenvironmental Engineering 126 (11), 969–978.

Sturm, J. F., 2002. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. Optimization Methods and Software 17 (6), 1105–1154.

The MathWorks, Inc, 2018. Optimization Toolbox^{TM} User's Guide. The MathWorks, Inc.

Vicente da Silva, M., Antão, A. N., 2008. Upper bound limit analysis with a parallel mixed finite element formulation. International Journal of Solids and Structures 45 (22), 5788–5804.

Widuliński, Ł., Tejchman, J., Kozicki, J., Leśniewska, D., 2011. Discrete simulations of shear zone patterning in sand in earth pressure problems of a retaining wall. International Journal of Solids and Structures 48 (7), 1191 – 1209.

Wörden, F. T., Achmus, M., 2013. Numerical modeling of three-dimensional active earth pressure acting on rigid walls. Computers and Geotechnics 51 (Supplement C), 83 – 90.