INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2010; **00**:1–25
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme

Examining the smooth and nonsmooth discrete element approaches to granular matter

M. Servin [1*] D. Wang[1], C. Lacoursière [1] and K. Bodin [1]

[1] *UMIT Research Lab, Umeå University, SE-90187, SWEDEN.*

SUMMARY

The smooth and nonsmooth approaches to the discrete element method (DEM) are examined from a computational perspective. The main difference can be understood as using explicit versus implicit time integration. A formula is obtained for estimating the computational effort depending on error tolerance, system geometric shape and size, and on the dynamic state. For the nonsmooth DEM, a regularized version mapping to the Hertz contact law is presented. This method has the conventional nonsmooth and smooth DEM as special cases depending on size of time step and value of regularization. The use of the projected Gauss-Seidel solver for nonsmooth DEM simulation is studied on a range of test systems. The following characteristics are found. Firstly, the smooth DEM is computationally more efficient for soft materials, wide and tall systems, and with increasing flow rate. Secondly, the nonsmooth DEM is more beneficial for stiff materials, shallow systems, static or slow flow and with increasing error tolerance. Furthermore, it is found that just as pressure saturates with depth in a granular column, due to force arching, also the required number of iterations saturates and become independent of system size. This effect make the projected Gauss-Seidel solver scale much better than previously thought. Copyright © 2010 John Wiley & Sons, Ltd.

Received . . .

KEY WORDS: discrete element method; multibody dynamics; granular media; contact; explicit time integration; linear solvers

1. INTRODUCTION

The discrete element method (DEM) of simulation is an important tool for studying many natural phenomena and structures in the fields of material science, statistical physics and geophysics when granular material is involved. It is also an important tool for design and optimization in the industries of processing, manufacturing, storage and transportation of granular materials, e.g., grains, minerals, pharmaceutical pills, pellets, sand and rocks. One of the main challenges for DEM simulation is to reduce the computational time of large-scale simulations with $10^3 - 10^9$ elements. There are two main approaches to discrete element simulation, referred to as *smooth* DEM and *nonsmooth* DEM. The difference lies in whether the viscoelastic nature of the contacts is resolved in time or not. The nonsmooth approach considers collisions and stick-slip frictional transitions as instantaneous events, where the velocity may change discontinuously in time, according to a given contact law. This allows for large time step integration with the potential of considerably less computational effort. Surprisingly few comparisons between smooth and nonsmooth DEM can be found in literature. One comparison of the computational scaling is made by Brendel *et al*, summarized by Eq. (1.37) and Fig. 6 in Ref. [4]. According to that analysis the nonsmooth approach is most favorable for dense,

*Correspondence to: Martin Servin, Department of Physics, Umeå University, SE-90187, SWEDEN. E-mail: martin.servin@physics.umu.se

quasistatic systems with few particles and high ratio of material stiffness over element mass while the smooth approach becomes increasingly favorable with increasing number of particles, kinetic energy and decreasing material stiffness. The comparison in Ref. [4] is, however, overly simplified and does not fully account for the dynamic state or geometric shape of the system and does not include any error tolerance threshold.

The main purpose of this paper is to provide a theoretical basis and empirical data for deciding when to use smooth or nonsmooth DEM simulation. Also, a regularized version of nonsmooth DEM is presented, referred to as *semi-smooth* DEM. It has the smooth and nonsmooth methods as limiting cases and thus constitutes a hybrid method sharing properties of both the other methods and forming a starting point for developing DEM simulation with adaptive transitions between smooth and nonsmooth time integration.

The paper is outlined as follows. In Sec. 2 an overview of smooth and nonsmooth DEM is given with key references. A hybrid version of the smooth and nonsmooth DEM is presented in Sec. 2.4. The computational properties of smooth and nonsmooth DEM are outlined and discussed in Sec. 3. The metrics used for comparison are introduced in Sec. 4.1. In Sec. 4 the simulation procedure and chosen test systems are described. The results of the simulations are presented and discussed in Sec. 5. The main conclusions are finally presented in Sec. 6.

1.1. Notation

For a rigid body, a, we use the notations $\vec{\boldsymbol{x}}_{[a]}, \vec{\boldsymbol{v}}_{[a]}, \vec{\boldsymbol{f}}_{[a]}$ and $m_{[a]}$ for position, velocity, force and mass, and $\vec{\boldsymbol{e}}_{[a]}, \vec{\boldsymbol{\omega}}_{[a]}, \vec{\boldsymbol{\tau}}_{[a]}$ and $\boldsymbol{I}_{[a]}$, for orientation, angular velocity, torque and inertia tensor. We agglomerate into generalized position, velocity, force and mass, denoted $\mathbf{x}_{[a]}, \mathbf{v}_{[a]}, \mathbf{f}_{[a]}$ and $\mathbf{M}_{[a]}$, with $\mathbf{v}_{[a]} = (\vec{\boldsymbol{v}}_{[a]}^{\mathrm{T}}, \vec{\boldsymbol{\omega}}_{[a]}^{\mathrm{T}})^{\mathrm{T}}$ etc. and $\mathbf{M}_{[a]} = \mathrm{diag}(m_{[a]}\mathbf{1}_{3\times 3}, \boldsymbol{I}_{[a]})$. These are components of the global system quantities that we denote $\mathbf{x}, \mathbf{v}, \mathbf{f}$ and $\mathbf{M}$. We use quaternions for representing orientations. The theory and algorithms in this paper applies to rigid bodies of general shape, although the simulations are made for spherically shaped rigid bodies, where integration of the orientation can be omitted. These bodies are referred to as particles but do posses rotational degree of freedom. Contact forces and velocities are sometimes decomposed in the directions of contact normals, $\vec{\mathbf{n}}$ and tangents, $\vec{\mathbf{t}}$. The gap function $g(\mathbf{x})$ measures the magnitude of overlap between contacting bodies. N_{p} denotes the number of particles and N_{c} the number of contacts. The matrix dimension of the global quantities are $\dim(\mathbf{x}) = 7N_{\mathrm{p}} \times 1$, $\dim(\mathbf{v}) = \dim(\mathbf{f}) = 6N_{\mathrm{p}} \times 1$, $\dim(\mathbf{M}) = 6N_{\mathrm{p}} \times 6N_{\mathrm{p}}$ and $\dim(\mathbf{g}) = N_{\mathrm{c}} \times 1$. Time is discretized in steps of size h and we denote the discrete time points by integer index i such that $t_i = t_{i-1} + h$. We use integer k for solver iteration step. Particle indices are emphasized by square brackets $[a]$ and $[b]$ and contact index by round brackets (α). The relative velocity at a contact point α between body a and b can thus be written $\vec{\boldsymbol{u}}^{(\alpha)} = \vec{\boldsymbol{v}}_{[a]} + \vec{\boldsymbol{d}}_{[a]}^{(\alpha)} \times \vec{\boldsymbol{\omega}}_{[a]} - \vec{\boldsymbol{v}}_{[b]} - \vec{\boldsymbol{d}}_{[b]}^{(\alpha)} \times \vec{\boldsymbol{\omega}}_{[a]}$, where $\vec{\boldsymbol{d}}_{[a]}^{(\alpha)}$ is the position of the contact point relative to $\vec{\boldsymbol{x}}_{[a]}$ and $\vec{\boldsymbol{d}}_{[b]}^{(\alpha)}$ relative to $\vec{\boldsymbol{x}}_{[b]}$. Furthermore, the notation $\mathbf{G}(n_1 : n_2, m_1 : m_2)$ is the standard Matlab notation for submatrices, referring to the rectangular submatrix block of $\mathbf{G}$ that ranges from row-column index (n_1, m_1) to (n_2, m_2).

2. SMOOTH AND NONSMOOTH DEM

An introduction to the theory and computational aspects of smooth DEM can be found in Ref. [24]. Comprehensive descriptions of nonsmooth DEM and its relation to smooth DEM can be found in Ref. [14] and in Ref. [25]. Important original works include those of Cundall and Strack [7] and Moreau [21]. In Sec. 2.4 we introduce a *semi-smooth* DEM that share features of both methods.

From a computational perspective, the main difference between smooth and nonsmooth DEM is related to explicit and implicit integration. In smooth DEM, the contact forces are modeled as damped springs or more general penalty functions [34, 22]. The equations of motion are the Newton-Euler equations of rigid body motion, which form a set of *ordinary differential equations* (ODE) and are usually integrated with an explicit time stepper using small step size depending on spring stiffness and particle mass. For an overview of explicit time stepping schemes, see [29].

DOI: 10.1002/nme

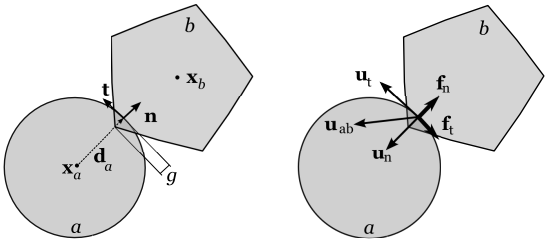

Figure 1. Overlapping discrete elements with notations.

Forces are computed *locally* for each contact pair. In nonsmooth DEM, the Newton-Euler equations are constrained by the Signorini-Coulomb contact law [14, 25], replacing the contact springs. The velocities are no longer assumed time-continuous. The contact forces and resulting velocities are computed *globally* by solving the constrained equations of motion for the entire contact network using an implicit time-integration algorithm. The nonsmooth method is a direct consequence of implicit time-integration in the presence of inequality conditions, e.g. that the contact force should vanish at separations or be limited by the Coulomb friction cone. As shown in Sec. 2.1, it is convenient to extend the system with auxiliary variables, *Lagrange multipliers* [34], and explicit use of constraints. Mathematically, this transforms the system from an ODE to a *differential variational inequality* (DVI) [23].

2.1. Relations between implicit and explicit integrators and constrained systems

In this section we elucidate the relations between implicit and explicit integrators, constrained systems and nonsmooth dynamics. Consider a point particle of unit mass in one dimension subject to the potential $U = \frac{1}{2\epsilon}x^2$, yielding the force $f = -\partial U/\partial x = -\epsilon^{-1}x$. Writing $v \equiv \dot{x}$, an explicit integration using the Verlet method with time step h yields

$$\begin{aligned} v_{i+1} &= v_i + hf_i = v_i - h\varepsilon^{-1}x_i \\ x_{i+1} &= x_i + hv_{i+1}. \end{aligned} \tag{1}$$

Here, only the force at time i is needed to advance the system. Using the implicit midpoint method, we instead have

$$\begin{aligned} v_{i+1} &= v_i + hf\left(\tfrac{x_{i+1}+x_i}{2}\right) = v_i - \tfrac{h}{2\varepsilon}(x_{i+1}+x_i) \\ x_{i+1} &= x_i + \tfrac{h}{2}(v_{i+1}+v_i). \end{aligned} \tag{2}$$

Simple manipulations yield

$$\left[1 + \tfrac{h^2}{4\varepsilon}\right] v_{i+1} = \left[1 - \tfrac{h^2}{4\varepsilon}\right] v_i - \tfrac{h}{\varepsilon}x_i. \tag{3}$$

This is a standard implicit integration and if the system was of higher dimension, the terms inside the brackets would be matrices and we would have to solve a linear system of equations for v_{i+1}. The terms containing ε^{-1} are problematic in the limit $\varepsilon \to 0$. Now introduce the auxiliary variable $\lambda = \varepsilon^{-1}x$ in continuous time so that, in discretized time,

$$\lambda = -\frac{1}{2\varepsilon}(x_{i+1}+x_i) = -\frac{1}{\varepsilon}x_i - \frac{h}{4\varepsilon}v_i - \tfrac{h}{4\varepsilon}v_{i+1}. \tag{4}$$

After simple manipulations, the stepping scheme reads

$$\begin{aligned} \begin{bmatrix} 1 & -1 \\ 1 & \frac{4\varepsilon}{h^2} \end{bmatrix} \begin{bmatrix} v_{i+1} \\ h\lambda \end{bmatrix} &= \begin{bmatrix} v_i \\ -\frac{4}{h}x_i - v_i \end{bmatrix} \\ x_{i+1} &= x_i + \frac{h}{2}(v_{i+1}+v_i). \end{aligned} \tag{5}$$

DOI: 10.1002/nme

There are no longer problematic terms of $1/\varepsilon$ and we can in principle set $\varepsilon = 0$. If the spring is one sided, i.e., inactive when $x > 0$ the integration should respect the condition $f_i = 0$ if $x_i > 0$.

If we now consider a more general mechanical system with mass matrix $\mathbf{M}$, coordinates $\mathbf{x}$ and velocity $\mathbf{v}$ subject to a strong potential $U = \frac{1}{2\varepsilon}\mathbf{g}^{\mathrm{T}}\mathbf{g}$, the force is then $\mathbf{f} = -\partial U/\partial \mathbf{x}^{\mathrm{T}} = -\frac{1}{\varepsilon}\mathbf{G}^{\mathrm{T}}\mathbf{g}$, where $\mathbf{G} = \partial \mathbf{g}/\partial \mathbf{x}$. Though the mass matrix $\mathbf{M}$ depends on the coordinates $\mathbf{x}$ in general, as in the case of the rigid body, this is neglected for simplicity. Following the same steps as before, the implicit integration can be written as either

$$\left[\mathbf{M} + \tfrac{h^2}{4\varepsilon}\mathbf{G}^{\mathrm{T}}\mathbf{G}\right]\mathbf{v}_{i+1} = \left[\mathbf{M} - \tfrac{h^2}{4\varepsilon}\mathbf{G}^{\mathrm{T}}\mathbf{G}\right]\mathbf{v}_i - \tfrac{h}{\varepsilon}\mathbf{G}^{\mathrm{T}}\mathbf{g}_i \tag{6}$$

or, by introducing auxiliary variable $\boldsymbol{\lambda} = -\varepsilon^{-1}\mathbf{g}(\mathbf{x})$

$$\begin{bmatrix} \mathbf{M} & -\mathbf{G}^{\mathrm{T}} \\ \mathbf{G} & \frac{4\varepsilon}{h^2} \end{bmatrix}\begin{bmatrix} \mathbf{v}_{i+1} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{v}_i + h\mathbf{f}_{\mathrm{s}} \\ -\frac{h}{4}\mathbf{g}_i - \mathbf{G}_i\mathbf{v}_i \end{bmatrix}, \tag{7}$$

where we absorbed the h factor in $\boldsymbol{\lambda}$ and added soft forces $\mathbf{f}_{\mathrm{s}}$ in view of what we will do below.

Clearly, if there were conditions on the forces $\boldsymbol{\lambda}$, they would have to be applied globally. That is not the case for the explicit method in which cases rules such as one-sided springs would be used, one force at a time at each step. The perturbation $4\varepsilon/(h^2)$ in the matrix in Eqn. (7) can be clearly related to a penalty. When it vanishes, however, we have a constrained system with hard contacts. From this perspective, the fundamental difference between smooth and nonsmooth DEM is related to explicit and implicit integration methods. The first one computes forces locally, the second globally.

The correspondence to constrained mechanical systems contains a number of subtleties with regards to the convergence of the trajectories, velocities, and forces λ in the limit where $\varepsilon \to 0$. The basic theory is simple however if we consider the Legendre transform of a potential $\mathbf{U} = \frac{1}{2\varepsilon}\mathbf{g}^{\mathrm{T}}\mathbf{g}$ where $\mathbf{g}(\mathbf{x})$ is a well behaved function of $\mathbf{x}$ such that the Jacobian matrix $\partial \mathbf{g}/\partial \mathbf{x} = \mathbf{G}$ has full rank at $\mathbf{g}(\mathbf{x}) = 0$. Choosing $\boldsymbol{\lambda}$ and the transform variable, the Legendre transform is defined as

$$\tilde{U}(\boldsymbol{\lambda}) = -\max_{\mathbf{g}}[\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{g} + \bar{U}(\mathbf{g})] \tag{8}$$

where $\bar{U}(\mathbf{g})$ is the pullback of $U(\mathbf{x})$. This leads to

$$\begin{aligned} U(\mathbf{x}) = \bar{U}(\mathbf{g}) = -\frac{\varepsilon}{2}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{g} \\ \varepsilon\boldsymbol{\lambda} + \mathbf{g} = 0. \end{aligned} \tag{9}$$

Note in particular that $\partial U/\partial \mathbf{x}^T = -\mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda}$. For a simple mechanical system with equations of motion $\mathbf{M}\ddot{\mathbf{x}} + \partial U/\partial \mathbf{x}^T - \mathbf{f}_{\mathrm{s}} = 0$, where $\mathbf{f}_{\mathrm{s}}$ are the "weak" forces, of the augmented equations of motion are then

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}} - \mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{f}_{\mathrm{s}} \\ \mathbf{g}(\mathbf{x}) + \varepsilon\boldsymbol{\lambda} = 0. \end{aligned} \tag{10}$$

It is possible to show that Eqn. (10) has a well behaved limit as $\varepsilon \downarrow 0$ provided there is some dissipation force of the form $\mathbf{f}_d = -\gamma\mathbf{G}^{\mathrm{T}}\dot{\mathbf{g}} = -\gamma\mathbf{G}^{\mathrm{T}}\mathbf{G}\dot{\mathbf{x}}$ [3]. At the limit $\varepsilon = 0$, we have the dynamics of a constrained system. Nonsmooth formulations are entirely based on constrained systems and this leads naturally to Differential Algebraic Equations (DAEs) of motion, or DVIs when including impacts and dry friction. Though the general numerical methods for these are computationally expensive [10], the special case of multibody systems allows simpler methods such as RATTLE[9].

In principle therefore, if we accept a nonzero relaxation, ε, there is no difference between the smooth (local) and nonsmooth (global) formulation other than the time integration method until inequalities are considered. However, it is clear from Eqn. (1) that inequalities and complementarity conditions, such as $0 \leq \mathbf{f} \perp \mathbf{g}(\mathbf{x}) \geq 0$, can be treated one at a time in the smooth formulation, but this is not the case when considering the system in Eqn. (7).

DOI: 10.1002/nme

2.2. *Smooth DEM*

In smooth DEM (SDEM) the contact normal force is a direct function of the geometric overlap function, $g(\mathbf{x})$, and it's time derivative. An example of a common force model is the nonlinear Hertz model that follows from the theory of linear viscoelastic materials. Friction is usually modeled as a spring in the tangential direction. The tangential spring extension is computed by integrating the slip velocity and the force is projected onto the friction cone to obey the Coulomb law. For spherical particles, the normal and tangential contact force are [24]

$$\vec{\mathbf{f}}_{\mathrm{n}} = k_{\mathrm{n}} \left(g^{3/2} + c g^{1/2} \dot{g} \right) \vec{\mathbf{n}} \tag{11}$$

$$\vec{\mathbf{f}}_{\mathrm{t}} = \mathrm{proj}_{\mu|\vec{\mathbf{f}}_{\mathrm{n}}|} \left(- \int k_{\mathrm{t}} \vec{\mathbf{u}}_{\mathrm{t}} \mathrm{d}t \right) \tag{12}$$

where $\vec{\mathbf{u}}_{\mathrm{t}}$ is the tangential relative velocity at the contact point. From the Hertz model for spheres the normal spring stiffness and damping coefficients are $k_{\mathrm{n}} = \mathrm{E}\sqrt{2d}/3(1-\nu^2)$ and $c = 4(1-\nu^2)(1-2\nu)\eta/15\mathrm{E}\nu^2$, where E is the Youngs's modulus, ν is the Poisson ratio and $d = (d_{[a]}^{-1} + d_{[b]}^{-1})^{-1}$ is the effective diameter from the interaction between spheres with diameter $d_{[a]}$ and $d_{[b]}$ and the material viscosity constant η [5]. For the friction spring coefficient k_{t} there is no such relation to fundamental material parameters and it must be determined by comparison with experimental results.

For explicit time integration methods, $(\mathbf{x}_i, \mathbf{v}_i) \rightarrow (\mathbf{x}_{i+1}, \mathbf{v}_{i+1})$, such as Verlet or symplectic Euler, the time step h must be smaller than the shortest time scale given by the viscoelastic interaction time $\sqrt{m/k_{\mathrm{n}}}$.

2.3. *Nonsmooth DEM*

In nonsmooth DEM (NDEM) the dynamics at short time scales is not resolved. Instead, the velocity is allowed to be discontinuous in time and contact forces, represented by impulses or kinematic constraints, can propagate through the system instantly. This enables large time step integration despite stiff materials. The contact force laws may be derived from the same penalty potentials as for smooth DEM or simply be determined directly empirically. We denote the impulse by $\vec{\mathbf{r}}$, instantly changing the contact velocity from $\vec{\mathbf{u}}^-$ to $\vec{\mathbf{u}}^+ = \vec{\mathbf{u}}^- + \mathbf{W}\vec{\mathbf{r}}$, with a transfer matrix $\mathbf{W}$ consistent with the preservation of total momentum. Over one time step h the time-averaged contact force is $\vec{\mathbf{f}} = \vec{\mathbf{r}}/h$. Specifically, the explicit contact force model in Eq. (11) and (12) is replaced by the *Signorini-Coulomb law* [14, 25] that if $g \geq 0$ then $\mathrm{law}_{\mathrm{SC}}[\vec{\mathbf{u}}_{i+1}, \vec{\mathbf{r}}_{i+1}] = \text{true}$:

$$0 \leq \vec{\mathbf{u}}_{\mathrm{n}} \perp \vec{\mathbf{f}}_{\mathrm{n}} \geq 0 \tag{13}$$

$$\vec{\mathbf{u}}_{\mathrm{t}} = 0 \quad \Longrightarrow \quad |\vec{\mathbf{f}}_{\mathrm{t}}| \leq \mu |\vec{\mathbf{f}}_{\mathrm{n}}| \tag{14}$$

$$\vec{\mathbf{u}}_{\mathrm{t}} \neq 0 \quad \Longrightarrow \quad |\vec{\mathbf{f}}_{\mathrm{t}}| = \mu |\vec{\mathbf{f}}_{\mathrm{n}}| \ , \ \ \vec{\mathbf{f}}_{\mathrm{t}}^{T} \vec{\mathbf{u}}_{\mathrm{t}} = -|\vec{\mathbf{f}}_{\mathrm{t}}||\vec{\mathbf{u}}_{\mathrm{t}}| \tag{15}$$

where the last condition is the maximum dissipation principle. We use $\mathrm{law}_{\mathrm{SC}}[\vec{\mathbf{u}}_{i+1}, \vec{\mathbf{r}}_{i+1}] = \text{true}$ in short for the inequalities (13)-(15). Newton's impact law of restitution, $\vec{\mathbf{u}}_{\mathrm{n}}^+ = -e\vec{\mathbf{u}}_{\mathrm{n}}^-$ with coefficient of restitution $0 \leq e \leq 1$, is conventionally also included in the contact law. The occurrence of time-discontinuities mediated by constraints or impulses changes the dynamic system from an ODE into a DVI. Solving the set of coupled inequalities (13)-(15) and update equations $(\mathbf{x}_i, \mathbf{v}_i) \rightarrow (\mathbf{x}_{i+1}, \mathbf{v}_{i+1})$ for the entire granular system is the main computational task in NDEM simulations, while it is a fast straight forward step in SDEM simulations. Observe that in NDEM, the magnitude of contact forces do not necessarily depend on the magnitude of the overlaps. Also, a local change in the system may propagate instantly through the entire contact network thanks to the implicit nature of the method. This produces the required bulk behaviour without need for resolving fast pressure waves. Observe that instant propagation speeds assumes perfectly rigid bodies but also for stiff materials the propagation length in one time-step may exceed the size of the system.

There are several methods for numerical integration of DVI systems. The approaches may be divided into iterative (splitting) solvers and direct (pivoting) solvers, but it is also possible to

DOI: 10.1002/nme

construct hybrid direct-iterative solvers. Often, the time-integration is made with a single-step first order method such as the symplectic (semi-implicit) Euler, resulting in a scheme nearly identical to the RATTLE algorithm [9] for molecular dynamics. In general, because the system is non-smooth there is no benefit from higher order methods, e.g., multistage or multistep. In Sec. 2.4 we present an extension of nonsmooth DEM that give the contact constraint force viscoelastic properties and is based on the linear complementarity formulation of nonsmooth DEM. The simulation result in the paper are produced with an implementation of this method. Therefore the linear complementarity formulation is outlined in more detail in Sec. 2.3.2.

2.3.1. Projected Gauss-Seidel. A natural and common approach is to treat each single contact problem, $\alpha, \beta = 1, 2, \ldots, N_\mathrm{c}$, sequentially and then iterate, $k = 1, 2, \ldots, N_\mathrm{it}$, until all contact laws are fulfilled to a desired error tolerance. A common algorithm for this is the blocked *projected Gauss-Seidel* (pGS) solver [13]. This is a stationary iterative method for solving Eqs. (13)-(15) approximately. At the k:th iteration step, each pair of contact velocity and impulse $(\vec{\mathbf{u}}_k^{(\alpha)}, \vec{\mathbf{r}}_k^{(\alpha)})$ is solved for each local contact problem α

$$\vec{\mathbf{u}}_k^{(\alpha)} - \mathbf{W}_{(\alpha\alpha)}\vec{\mathbf{r}}_k^{(\alpha)} = \vec{\mathbf{u}}_\mathrm{s}^{(\alpha)} + \Delta\vec{\mathbf{u}}_k^{(\alpha)} \tag{16}$$

$$\mathrm{law}_\mathrm{SC}\left[\vec{\mathbf{u}}_k^{(\alpha)}, \vec{\mathbf{r}}_k^{(\alpha)}\right] = \mathrm{true} \tag{17}$$

where

$$\Delta\vec{\mathbf{u}}_k^{(\alpha)} = \sum_{\beta<\alpha} \mathbf{W}_{(\alpha\beta)}\vec{\mathbf{r}}_k^{(\beta)} + \sum_{\beta>\alpha} \mathbf{W}_{(\alpha\beta)}\vec{\mathbf{r}}_{k-1}^{(\beta)} \tag{18}$$

and $\vec{\mathbf{u}}_\mathrm{s}$ is the contact velocity as would be in the presence of only smooth forces, $\mathbf{W}_{(\alpha\beta)} = \mathbf{H}^\mathrm{T}_{(\alpha)[a]}\mathbf{M}^{-1}_{[ab]}\mathbf{H}_{(\beta)[b]}$ is the Delassus operator for contact α and β with $\mathbf{H}^\mathrm{T}_{(\alpha)[a]}$ being the affine transformation of velocity of body a to relative contact velocity in point α and $\mathbf{M}$ the mass matrix of the system. After convergence to a set of impulses and contact velocities fulfilling law_SC to desired tolerance, the body velocities in global coordinates are computed by $\mathbf{v}_{i+1} = \mathbf{v}_\mathrm{s} + \mathbf{H}\mathbf{r}$ with the final net impulse $\mathbf{H}\mathbf{r}$ and $\mathbf{v}_\mathrm{s} = \mathbf{v}_i + h\mathbf{M}^{-1}\mathbf{f}_\mathrm{s}$ is the updated velocity in presence of smooth forces. Finally, position is updated by $\mathbf{x}_{i+1} = \mathbf{x}_i + h\mathbf{v}_{i+1}$. In $\theta = 1/2$ integration methods, the positions are first update on half time step with the velocities from smooth forces before contacts are computed.

2.3.2. Mixed linear complementarity problem. Another popular method follows from linearization of the Coulomb friction law, by approximating the friction cone with a (scaled) box or a polyhedral cone. The constrained equations of motion may then be put in the form of a *mixed linear complementarity problem* (MLCP). In terms of body velocity $\mathbf{v}$ and Lagrange multiplier $\boldsymbol{\lambda}$ the MLCP reads [16]

$$\begin{gathered}\mathbf{H}\mathbf{z} + \mathbf{b} = \mathbf{w}_+ - \mathbf{w}_- \\ 0 \le \mathbf{z} - \mathbf{l} \perp \mathbf{w}_+ \ge 0 \\ 0 \le \mathbf{u} - \mathbf{z} \perp \mathbf{w}_- \ge 0\end{gathered} \tag{19}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{M} & -\mathbf{G}_\mathrm{n}^\mathrm{T} & -\mathbf{G}_\mathrm{t}^\mathrm{T} \\ \mathbf{G}_\mathrm{n} & 0 & 0 \\ \mathbf{G}_\mathrm{t} & 0 & 0 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{v}_{i+1} \\ \boldsymbol{\lambda}_{\mathrm{n},i+1} \\ \boldsymbol{\lambda}_{\mathrm{t},i+1} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -\mathbf{M}\mathbf{v}_\mathrm{s} \\ 0 \\ 0 \end{bmatrix} \tag{20}$$

$\mathbf{G}_\mathrm{n}$ and $\mathbf{G}_\mathrm{t}$ are the normal and tangential constraint Jacobians for the Signorini-Coulomb law and $\mathbf{w}_\pm$ are (temporary) slack variables, $\mathbf{l}$ and $\mathbf{u}$ are the upper and lower limits on the solution implied by the linearized Signorini-Coulomb law. The normal and tangential contact constraint force is $\mathbf{f}_\mathrm{n} = \mathbf{G}_\mathrm{n}^\mathrm{T}\boldsymbol{\lambda}_\mathrm{n}/h$ and $\mathbf{f}_\mathrm{t} = \mathbf{G}_\mathrm{t}^\mathrm{T}\boldsymbol{\lambda}_\mathrm{t}/h$, respectively. The h^{-1} factor is due to a convenient normalization of the Lagrange multiplier such that it has the dimension of impulse.

The MLCP formulation with saddle-point matrix $\mathbf{H}$ in Eq. (20) is common in the realm of *multibody system dynamics* for modeling of linked mechanism using kinematic constraints for describing various joints, their actuations and geometric limits. One significant advantage of this

DOI: 10.1002/nme

formulation of nonsmooth DEM is therefore that it automatically provides a unified formulation for modeling and simulating of granular material strongly coupled with mechatronic systems, such as vehicles, conveying systems and articulated manipulators [18].

A generalization of the MLCP formulation to include also viscoelastic properties in the contact dynamics is presented in Sec. 2.4 and used in simulation in Sec. 4. Therefore we provide some further details on the MLCP formulation here. The normal constraint force acts to prevent penetration, i.e., to maintain the constraint $g^{(\alpha)} \leq 0$, for each contact candidate α between body a and b. The overlap is computed

$$g^{(\alpha)} = \vec{\mathbf{n}}_{(\alpha)}^{\mathrm{T}}(\vec{\mathbf{x}}_{[a]} + \vec{\mathbf{d}}_{[a]}^{(\alpha)} - \vec{\mathbf{x}}_{[b]} - \vec{\mathbf{d}}_{[b]}^{(\alpha)}) \tag{21}$$

with the contact normal $\vec{\mathbf{n}}_{(\alpha)}$ directed outwards from a, and $\vec{\mathbf{d}}_{[a]}^{(\alpha)}$ is the position of the contacting point on the surface of body a relative to its center of mass position $\vec{\mathbf{x}}_{[a]}$, see Fig. 1. The nonpenetration constraint implies that the relative contact normal velocity should be zero or separating, i.e., $\mathbf{G}_{\mathrm{n}}^{(\alpha)}\mathbf{v} \geq 0$, with the normal Jacobian $\mathbf{G}_{\mathrm{n}}^{(\alpha)} = \partial g^{(\alpha)}/\partial \mathbf{x}$. The nonzero blocks of the normal Jacobian are

$$\mathbf{G}_{\mathrm{n}[a]}^{(\alpha)} = \begin{bmatrix} -\vec{\mathbf{n}}_{(\alpha)}^{\mathrm{T}} & -(\vec{\mathbf{d}}_{[a]}^{(\alpha)} \times \vec{\mathbf{n}}_{(\alpha)})^{\mathrm{T}} \end{bmatrix} \tag{22}$$

$$\mathbf{G}_{\mathrm{n}[b]}^{(\alpha)} = \begin{bmatrix} \vec{\mathbf{n}}_{(\alpha)}^{\mathrm{T}} & (\vec{\mathbf{d}}_{[b]}^{(\alpha)} \times \vec{\mathbf{n}}_{(\alpha)})^{\mathrm{T}} \end{bmatrix} \tag{23}$$

Friction is introduced as a constraint of vanishing relative contact velocity $\mathbf{v}_{[ab]}^{(\alpha)}$ in the tangent plane, i.e., $\mathbf{G}_{\mathrm{t}}^{(\alpha)}\mathbf{v} = 0$, unless the force reach the friction bounds. With the *box friction* approximation the tangent plane is spanned with two orthogonal vectors $\mathbf{t}_1^{(\alpha)\mathrm{T}}$ and $\mathbf{t}_2^{(\alpha)\mathrm{T}}$ and each friction multiplier has two components $\boldsymbol{\lambda}_{\mathrm{t}}^{(\alpha)} = [\lambda_{\mathrm{t}1}^{(\alpha)}\ \lambda_{\mathrm{t}2}^{(\alpha)}]^{\mathrm{T}}$. The nonzero blocks of the tangent Jacobian are

$$\mathbf{G}_{\mathrm{t}[a]}^{(\alpha)} = \begin{bmatrix} -\vec{\mathbf{t}}_1^{(\alpha)\mathrm{T}} & -(\vec{\mathbf{d}}_{[a]}^{(\alpha)} \times \vec{\mathbf{t}}_1^{(\alpha)})^{\mathrm{T}} \\ -\vec{\mathbf{t}}_2^{(\alpha)\mathrm{T}} & -(\vec{\mathbf{d}}_{[a]}^{(\alpha)} \times \vec{\mathbf{t}}_2^{(\alpha)})^{\mathrm{T}} \end{bmatrix} \tag{24}$$

$$\mathbf{G}_{\mathrm{t}[b]}^{(\alpha)} = \begin{bmatrix} \vec{\mathbf{t}}_1^{(\alpha)\mathrm{T}} & (\vec{\mathbf{d}}_{[b]}^{(\alpha)} \times \vec{\mathbf{t}}_1^{(\alpha)})^{\mathrm{T}} \\ \vec{\mathbf{t}}_2^{(\alpha)\mathrm{T}} & (\vec{\mathbf{d}}_{[b]}^{(\alpha)} \times \vec{\mathbf{t}}_2^{(\alpha)})^{\mathrm{T}} \end{bmatrix} \tag{25}$$

The Jacobian blocks have dimension $\dim(\mathbf{G}_{\mathrm{n}[a]}^{(\alpha)}) = 1 \times 6$ and $\dim(\mathbf{G}_{\mathrm{t}[a]}^{(\alpha)}) = 2 \times 6$. The assembled constraint vector $\mathbf{g} = [g^{(1)}\ g^{(2)}\ \ldots\ g^{(N_{\mathrm{c}})}]^{\mathrm{T}}$ has dimension $N_{\mathrm{c}} \times 1$, the assembled Jacobians $\dim(\mathbf{G}_{\mathrm{n}}) = N_{\mathrm{c}} \times 6N_{\mathrm{p}}$ and $\dim(\mathbf{G}_{\mathrm{t}}) = 2N_{\mathrm{c}} \times 6N_{\mathrm{p}}$, and the Lagrange multipliers $\dim(\boldsymbol{\lambda}_{\mathrm{n}}) = N_{\mathrm{c}} \times 1$ and $\dim(\boldsymbol{\lambda}_{\mathrm{t}}) = 2N_{\mathrm{c}} \times 1$. The resulting MLCP thus consist of a sparse saddle-point matrix of size $\dim(\mathbf{H}) = (6N_{\mathrm{p}} + 3N_{\mathrm{c}}) \times (6N_{\mathrm{p}} + 3N_{\mathrm{c}})$ and $\dim(\mathbf{z}) = 6N_{\mathrm{p}} + 3N_{\mathrm{c}}$ variables. Improving the approximation of the friction cone from a box to a polyhedron result in similar system but with more auxiliary variables.

Contacts are separated into *continuous contacts* and *impacting contacts*. Impacting contacts are those not occurring with last time step. The effect of impacts are treated in an *impact stage*, solving a MLCP based on the Newton impact law $\mathbf{G}_{\mathrm{n}}\mathbf{v}_{[ab]}^{+} = -e\mathbf{G}_{\mathrm{n}}\mathbf{v}_{[ab]}^{-}$ with restitution coefficient e, before proceeding wth the main solve and time integration. See, Appendix A for further details.

There are three type of solvers for MLCPs: pivoting methods, Newton (line search) methods and iterative methods with different requirements on the matrix $\mathbf{H}$, e.g., being positive-definite. Regularization is needed for handling ill-posed or ill-conditioned problems, e.g., due to existence multiple solutions to the contact problem, contact constraint degeneracy and occurrence of large mass ratios. Regularization terms are entered as positive diagonal perturbations in $\mathbf{H}$ and stabilization terms on the form $\alpha\mathbf{g} + \beta\mathbf{G}_{\mathrm{n}}\mathbf{v}$ are introduced in the $\mathbf{b}$ vector which are required for restoring constraint violations if they should occur. These terms bring solvability and numerical stability and corresponds to a solving a slightly different physical system than the original one. The systems converge to (12)-(14) weakly in the limit of zero regularization. See, e.g., Ref. [17].

DOI: 10.1002/nme

2.4. Semi-smooth DEM

In this section we present a *semi-smooth DEM* that shares the features of large time step integration with nonsmooth DEM and the viscoelastic interaction forces of smooth DEM. Elastic contact models for nonsmooth DEM by regularization corresponding to linear springs was recently presented and examined in Ref. [15]. In the present paper we describe a regularization that maps to nonlinear springs according to the Hertz contact law. This is realized by considering the nonsmooth DEM as the stiff limit of smooth DEM with particular energy potentials and dissipation functions, for generating the constraint regularization and stabilization terms to the MLCP form in Sec. 2.3.2. The theoretical basis is thoroughly described in Ref. [17] and has previously been applied to large time step simulation of large-scale granular flows for geometric design of pelletizing drum [32].

Firstly, observe that the normal contact force in Eq. (11) follows from $\mathbf{f}_\mathrm{n} = -\nabla_x U_\mathrm{n} - \nabla_v \mathcal{R}_\mathrm{n}$ with the following potential energy function and Rayleigh dissipation functions

$$U_\mathrm{n} = \frac{1}{2\varepsilon_\mathrm{n}} \bar{\mathbf{g}}^\mathrm{T} \bar{\mathbf{g}} \tag{26}$$

$$\mathcal{R}_\mathrm{n} = \frac{1}{2\gamma_\mathrm{n}} (\bar{\mathbf{G}}_\mathrm{n}\mathbf{v})^\mathrm{T} (\bar{\mathbf{G}}_\mathrm{n}\mathbf{v}) \tag{27}$$

with $\bar{g}_{(\alpha)} = g_{(\alpha)}^{e_\mathrm{H}}$, $\bar{\mathbf{G}}_\mathrm{n}^{(\alpha)} = e_\mathrm{H} g_{(\alpha)}^{e_\mathrm{H}-1} \mathbf{G}_\mathrm{n}^{(\alpha)}$, $\varepsilon_\mathrm{n}^{-1} = k_\mathrm{n}/e_\mathrm{H}$, $\gamma_\mathrm{n}^{-1} = k_\mathrm{n} c/e_\mathrm{H}$ and exponent $e_\mathrm{H} = 5/4$ for the nonlinear Hertz model with stiffness and damping parametrized as in Sec. 2.2. Note the relation $\dot{\bar{\mathbf{g}}} = \bar{\boldsymbol{G}}\mathbf{v}$. Secondly, instead of treating contact forces explicitly as gradients of potentials these are introduced as constraint forces $\bar{\mathbf{G}}^\mathrm{T}\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}$ are introduced as auxiliary variables obeying $\varepsilon\boldsymbol{\lambda} + \bar{\mathbf{g}} + \tau\bar{\mathbf{G}}\mathbf{v} = 0$, with regularization and damping parameters ε and τ. Thirdly, employing a time-discrete formulation of the variational principle leads to the following first order fixed time step integration method, coined SPOOK [17],

$$\begin{array}{c} \bar{\mathbf{H}}\mathbf{z} + \bar{\mathbf{b}} = \mathbf{w}_+ - \mathbf{w}_- \\ 0 \le \mathbf{z} - \mathbf{l} \perp \mathbf{w}_+ \ge 0 \\ 0 \le \mathbf{u} - \mathbf{z} \perp \mathbf{w}_- \ge 0 \end{array} \tag{28}$$

where

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{M} & -\bar{\mathbf{G}}_\mathrm{n}^\mathrm{T} & -\bar{\mathbf{G}}_\mathrm{t}^\mathrm{T} \\ \bar{\mathbf{G}}_\mathrm{n} & \boldsymbol{\Sigma} & 0 \\ \bar{\mathbf{G}}_\mathrm{t} & 0 & \boldsymbol{\Gamma} \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{v}_{i+1} \\ \boldsymbol{\lambda}_{\mathrm{n},i+1} \\ \boldsymbol{\lambda}_{\mathrm{t},i+1} \end{bmatrix}, \quad \bar{\mathbf{b}} = \begin{bmatrix} -\mathbf{M}\mathbf{v}_\mathrm{s} \\ \frac{4}{h}\boldsymbol{\Upsilon}\bar{\mathbf{g}} - \boldsymbol{\Upsilon}\bar{\mathbf{G}}_\mathrm{n}\mathbf{v}_i \\ 0 \end{bmatrix} \tag{29}$$

and regularization and stabilization coefficients

$$\boldsymbol{\Sigma} = \frac{4}{h^2}\frac{\varepsilon_\mathrm{n}}{1+4\frac{\tau_\mathrm{n}}{h}}\mathbf{1}_{N_c\times N_c}, \quad \boldsymbol{\Gamma} = \frac{\gamma_\mathrm{t}}{h}\mathbf{1}_{2N_c\times 2N_c}, \quad \boldsymbol{\Upsilon} = \frac{1}{1+4\frac{\tau_\mathrm{n}}{h}}\mathbf{1}_{N_c\times N_c} \tag{30}$$

The parameter τ_n controls the dissipation rate in the normal force. We use $\tau_\mathrm{n} = \max(n_\mathrm{s}h, \gamma_\mathrm{n}^{-1})$ with $n_\mathrm{s} = 2$ and the clamping is enforced to guarantee numerical stability at large time step integration. Similarly, the Coulomb friction is modeled as the stiff limit of a Rayleigh dissipation function for the relative velocity in the contact tangent plane with and dissipation rate γ_t and limits on the tangential Lagrange multiplier and $\bar{\mathbf{G}}_\mathrm{t} = \mathbf{G}_\mathrm{t}$. An algorithm for the semi-smooth DEM with projected Gauss-Seidel is presented in Appendix B.

Both the smooth DEM and nonsmooth DEM can be recovered from the semi-smooth DEM. Nonsmooth DEM follows directly from taking the stiff limit $\varepsilon, \gamma \to 0$. This implies $\boldsymbol{\Gamma}, \boldsymbol{\Sigma}, \boldsymbol{\Upsilon} \to 0$, and Eq. (28)-(29) reduces exactly to Eq. (19)-(20). Smooth DEM, on the other hand, is recovered in the limit of time step $h \to 0$ with $\varepsilon, \gamma, \tau$ fix. Observe that in this limit

$$\boldsymbol{\Sigma} \to \frac{\varepsilon_\mathrm{n}}{\tau_\mathrm{n} h}\mathbf{1}_{N_c\times N_c}, \quad \boldsymbol{\Upsilon} \to \frac{h}{4\tau_\mathrm{n}}\mathbf{1}_{N_c\times N_c}, \quad \tfrac{4}{h}\boldsymbol{\Upsilon}\bar{\mathbf{g}} - \boldsymbol{\Upsilon}\bar{\mathbf{G}}_\mathrm{n}\mathbf{v}_i \to \tau_\mathrm{n}^{-1}\bar{\mathbf{g}} \tag{31}$$

Assuming smoothness in $\mathbf{v}$, it is possible to eliminate $\boldsymbol{\lambda}_\mathrm{n}$ and $\boldsymbol{\lambda}_\mathrm{t}$ in the MLCP in Eq. (28) whereby it simplifies to the stepping rule $\mathbf{M}\mathbf{v}_{i+1} = \mathbf{M}\mathbf{v}_{\mathrm{s},i+1} + \bar{\mathbf{G}}_\mathrm{n}^\mathrm{T}\boldsymbol{\lambda}_\mathrm{n} + \bar{\mathbf{G}}_\mathrm{t}^\mathrm{T}\boldsymbol{\lambda}_\mathrm{t}$ where

$$\bar{\mathbf{G}}_\mathrm{n}^{(\alpha)\mathrm{T}}\lambda_\mathrm{n}^{(\alpha)}/h \approx \bar{\mathbf{G}}_\mathrm{n}^{(\alpha)\mathrm{T}}\left(\tfrac{\tau h}{\varepsilon}\bar{\mathbf{G}}_\mathrm{n}^{(\alpha)}\mathbf{v}_{i+1} + \tfrac{\tau h}{\varepsilon\tau}\bar{g}_{(\alpha)}\right)/h \approx k_\mathrm{n}\left(g_{(\alpha)}^{3/2} + c g_{(\alpha)}^{1/2}\dot{g}_{(\alpha)}\right)\vec{\boldsymbol{n}}^{(\alpha)} \equiv \mathbf{f}_\mathrm{n}^{(\alpha)} \tag{32}$$

DOI: 10.1002/nme

$$\bar{\mathbf{G}}_{\mathrm{t}}^{(\alpha)\mathrm{T}}\boldsymbol{\lambda}_{\mathrm{t}}^{(\alpha)}/h \approx \mathrm{proj}_{\mu|\mathbf{f}_{\mathrm{n}}^{(\alpha)}|}\left(\bar{\mathbf{G}}_{\mathrm{t}}^{(\alpha)\mathrm{T}}(\gamma^{-1}\bar{\mathbf{G}}_{\mathrm{t}}^{(\alpha)}\mathbf{v}^{i})\right) \approx \mathrm{proj}_{\mu|\mathbf{f}_{\mathrm{n}}^{(\alpha)}|}\left(k_{\mathrm{t}}\vec{\boldsymbol{u}}_{\mathrm{t},i}^{(\alpha)}\right) \equiv \mathbf{f}_{\mathrm{t}}^{(\alpha)} \tag{33}$$

in which we have approximated $\bar{\mathbf{G}}_{\mathrm{n}}\mathbf{v}_{i+1} \approx \dot{\bar{\mathbf{g}}}(t_i)$ and $\bar{\mathbf{G}}_{\mathrm{t}}^{\mathrm{T}}\bar{\mathbf{G}}_{\mathrm{t}}\mathbf{v}_{i+1} \approx -\mathbf{u}_{\mathrm{t},i}$ and used that $k_{\mathrm{n}} = e_{\mathrm{H}}\varepsilon^{-1}$, $k_{\mathrm{t}} = \gamma_{\mathrm{t}}^{-1}$, $c = \tau e_{\mathrm{H}}$ and $e_{\mathrm{H}} = 5/4$. Hence, the multipliers can be removed by substitution and we obtain an identical time stepping scheme as for smooth DEM when applying a semi-implicit Euler discretization. We thus conclude that the results of semi-smooth DEM and smooth DEM coincide in the limit of small time steps.

The elastic properties are present also for large time steps. Specifically, in the quasi-static regime, the regularization corresponds directly to the Hertz model of elastic contacts. Assume $\bar{\mathbf{G}}_{\mathrm{n}}\mathbf{v}_{i+1} = \bar{\mathbf{G}}_{\mathrm{n}}\mathbf{v}_i \approx 0$ and Eq. (28)-(29) implies that $\boldsymbol{\Sigma}\boldsymbol{\lambda}_{\mathrm{n}} = (4/h)\boldsymbol{\Upsilon}\bar{\mathbf{g}}$ and the constraint force, $\mathbf{f}_{\mathrm{n}}^{(\alpha)} = \bar{\mathbf{G}}_{\mathrm{n[b]}}^{(\alpha)\mathrm{T}}(1,1:3)\lambda_{\mathrm{n}}^{(\alpha)}/h$ on body b from contact α with body a become

$$\mathbf{f}_{\mathrm{n}}^{(\alpha)} = \Sigma^{-1}(4/h^2)\Upsilon\bar{\mathbf{G}}_{\mathrm{n[b]}}^{\alpha\mathrm{T}}(1,1:3)\bar{g}_{(\alpha)} = k_{\mathrm{n}}g_{(\alpha)}^{3/2}\vec{\mathbf{n}}^{(\alpha)} \tag{34}$$

which equals the smooth normal force in Eq. (11) in the case of zero relative contact velocity.

3. COMPUTATIONAL PROPERTIES

In this section we make a theoretical comparison between of the computational properties of smooth and nonsmooth DEM. The semi-smooth DEM has identical computational complexity as the nonsmooth and is therefore not covered separately. In particular, we consider how the computational effort scale with the number of particles and dependency on the dynamical state and geometric shape of the systems. The computational time, τ_{DEM}, for simulating a process lasting for τ_{real} units of real time is

$$\tau_{\mathrm{DEM}} = \frac{\Omega_{\mathrm{DEM}}}{h_{\mathrm{DEM}}}\tau_{\mathrm{real}} \tag{35}$$

where Ω_{DEM} is the required computational time for advancing the simulation one time step of size h_{DEM}. The computational effort is $\tau_{\mathrm{DEM}}/\tau_{\mathrm{real}}$. The main steps of the simulation loop are *contact detection*, *solve* and *simulation management and I/O*. The contact detection step involves a *broad phase*, where contact candidates are found, and a *narrow phase*, where contact positions and overlap magnitudes are computed. The *solve phase* refer to the numeric integration of the equations of motion.

3.1. Smooth DEM

For smooth DEM with explicit or semi-implicit time integration, the solve stage consists of simple evaluation and summation of forces and update of velocities and positions involving only a few multiplication and additions per particle. The computational bottleneck lies in the contact detection stage. A typical smooth DEM simulation spends roughly 80% of the computational time on contact detection [33] and, assuming no parallelization, scales at best linearly with N_{p} and as worst as $N_{\mathrm{p}}\log(N_{\mathrm{p}})$ if temporal coherence cannot be exploited. For numerical stability the time step size is limited by $h_{\mathrm{SDEM}} \lesssim \sqrt{m/k}$ and the computational time is thus

$$\tau_{\mathrm{SDEM}} = \sqrt{\frac{k}{m}}K_{\mathrm{SDEM}}N_{\mathrm{p}}\tau_{\mathrm{real}} \tag{36}$$

where we have split $\Omega_{\mathrm{SDEM}} = K_{\mathrm{SDEM}}N_{\mathrm{p}}$ and K_{SDEM} is the average computational time per time step and particle. The factor K_{SDEM} depends on hardware, software, geometric shape of particles and collision detection algorithm and its implementation. With a conventional desktop computer (specified in Appendix C) running the smooth DEM software library LIGGGTHS [19] with spherical particles it was measured [11] $K_{\mathrm{SDEM}} \sim 10^{-6}$.

DOI: 10.1002/nme

3.2. Nonsmooth DEM

For nonsmooth DEM the solve stage dominates the computational time, e.g., 88% of the time was reported in [28]. Typically, the computational time for one time step is much larger than for smooth DEM. The benefit instead comes from integration with fewer and larger time steps.

The time step limit in nonsmooth DEM is set by the characteristic relative velocities or gravity acceleration. In one integration time step these should not cause an impact overlap larger than some fraction ϵ of the particle diameter, i.e., $h_{\text{NDEM}} \leq \epsilon d/v_{\text{n}}$ and $h_{\text{NDEM}} \leq \sqrt{2\epsilon d/g}$. This introduce ϵ as an error tolerance. The computational time become

$$\tau_{\text{NDEM}} = \frac{\Omega_{\text{NDEM}}}{\min\left(\epsilon\frac{d}{v_{\text{n}}}, \sqrt{\frac{2\epsilon d}{g}}\right)}\tau_{\text{real}} \tag{37}$$

Contrary to smooth DEM the computational time per time step Ω_{NDEM} is in general not a linear function of the number of particles (or number of contacts). The equations of motions for nonsmooth DEM are differential variational inequality (DVI) whose computational properties are largely open questions [6], lacking proof of existence and uniqueness of solutions as well as of general proof of convergence and numerical stability of most solution algorithms. Both theoretical and empirical analysis are complicated by the fact that the solutions are in general discontinuous with respect to change in initial data. The computational scaling depends ultimately on the choice of numerical solver as well as of the dynamical and geometric state of the system.

A theoretical upper limit of the computational efficiency can be found by considering the linear system $Hz + b = 0$ of the MLCP in Eq. (19) and ignoring the complementarity conditions. The matrix size is of the order $3N_{\text{c}} \times 3N_{\text{c}}$. The number of contacts is related to the number of particles as $N_{\text{c}} \sim (n_{\text{c}}/2)N_{\text{p}}$, assuming each particle on average has n_{c} neighbours, which has a number ranging between $2, 3-8, 6-12$ depending on the system dimensionality ranging between 1D (chain), 2D (plane) and 3D (bulk) and on the packing density. The matrix H is block-sparse and the linear system can in many cases be solved efficiently using direct or iterative methods exploiting the block-sparseness. The best theoretical scaling for direct solvers is provided by algorithms using *nested dissection* for reordering and factorization (the computationally most expensive part). This has time complexity of $\Omega_{\text{NDEM}} \sim \mathcal{O}(N_{\text{c}}^{(1+n_{\text{d}})/2})$, for $n_{\text{d}} = 1-3$ dimensionality. For the case of a 1D column of particles (chain) the computational time scales linearly with number of contacts and particles, $\Omega_{\text{NDEM}} \sim \mathcal{O}(N_{\text{c}})$, but for higher dimensionality the scaling is in general superlinear. Using the AgX Multiphysics Toolkit [1] we measured $\Omega_{\text{NDEM}} = K_{\text{NDEM}}N_{\text{c}}$ with $K_{\text{NDEM}} = 10^{-5}$ s for the 1D column using a block-sparse LU solver. When machine precision is not required an iterative solver may give the solution to a given tolerance faster. *Multigrid* methods and *conjugate gradient* (CG) [26, 27] method may scale almost linearly with N_{c} but many issues remain in transferring these results to irregular contact networks with complementarity conditions.

3.2.1. Efficiency of the Gauss-Seidel solver. The Gauss-Seidel algorithm (GS) is commonly considered a poor choice for solving linear systems. Still, it is of common use for integrating nonsmooth DEM simulations. Although the asymptotic convergence is slow[†] the initial convergence may be fast and the algorithm allows changes in the active set without restarting, as the CG does. The projected block Gauss-Seidel algorithm solves the local contact problem well but approaches to the global solution in a diffusive manner with increasing number of iterations. The residual from truncating at finite number of iterations result in *numeric elasticity* [30] with an effective sound velocity $v_{\text{GS}} = \sqrt{N_{\text{it}}}d/h_{\text{NDEM}}$. Consequently, for a system with side length l and particle size d, the required number of iterations for establishing the contact force network required for maintaining perfect rigidity scales as $N_{\text{it}} \propto (l/d)^2 \sim N_{\text{c}}^{2/n_{\text{d}}}$. Each iteration step involves solving each of the N_{c}

[†]For the Gauss-Seidel algorithm the residual decays asymptotically as the logarithm of the spectral radius of the iteration matrix.

DOI: 10.1002/nme

two-body contact problem. The resulting computational time complexity thus become

$$\Omega^{\text{GS}}_{NDEM} = K^{\text{GS}}_{NDEM} N_{\text{it}} N_{\text{c}} \tag{38}$$

where K^{GS}_{NDEM} is the average computational time for solving each two-body problem. Consequently, projected Gauss-Seidel scales $\Omega_{\text{NDEM}} \sim \mathcal{O}(N_{\text{c}}^{1+2/n_{\text{d}}})$. In our implementation, described in Appendix B, we measure the scaling coefficient to $K^{\text{GS}}_{NDEM} = 10^{-6}$ s with a conventional desktop computer (specified in Appendix C).

3.3. Comparing smooth and nonsmooth DEM

To compare the difference in computational efficiency between the smooth and nonsmooth DEM we consider the ratio of their computational time from Eq. (36) and (37)

$$\frac{\tau_{\text{NDEM}}}{\tau_{\text{SDEM}}} = \sqrt{\frac{\max\left(\epsilon^{-2} m v_{\text{n}}^2, 2\epsilon^{-1} mgd\right)}{kd^2}} \frac{\Omega_{\text{NDEM}}}{K_{\text{SDEM}} N_{\text{p}}} \tag{39}$$

When this ratio is smaller than one the nonsmooth DEM is more efficient and vice versa. The square root term is the effect of different time step size h_{SDEM} and h_{NDEM}. The relative efficiency of the nonsmooth DEM increase with particle size and stiffness and decrease with increasing relative velocity and mass. The second term on the right hand side is the effect of different computational scaling of the solvers for smooth and nonsmooth DEM.

Example. Consider particles of size $d = 0.01$ m, mass density 2500 kg/m, Young's modulus E $= 10$ GPa and average normal contact velocity $v_{\text{n}} \lesssim 0.1$ m/s. This means $m \sim 10^{-3}$ kg and $k \sim 10^9$. The resulting smooth and nonsmooth time step size become $h_{\text{SDEM}} \lesssim \sqrt{m/k} = 10^{-6}$ and $h_{\text{NDEM}} = 10^{-3}$ s for impact overlap tolerance set to $\epsilon = 0.01$. The time step ratio (the square root in Eq. (39)) become $h_{\text{NDEM}}/h_{\text{SDEM}} \approx 10^{-3}$. In the case of 1D dimensionality, e.g., a column of particles, and employing a direct solver for the nonsmooth DEM with $K_{\text{NDEM}} = 10^{-5}$ (measured from AgX) and a smooth DEM implementation with $K_{\text{SDEM}} \sim 10^{-6}$ (measured from LIGGGHTS) we obtain the time ratio $\frac{\tau_{\text{NDEM}}}{\tau_{\text{SDEM}}} \approx 10^{-2}$. Hence, for the column configuration the nonsmooth DEM is 100 times faster than the smooth DEM irrespective of the number of particles in the column. If we instead consider an iterative Gauss-Seidel solver, with scaling $K^{\text{GS}}_{NDEM} = 10^{-6}$ for the nonsmooth DEM we get the time ratio $\frac{\tau_{\text{NDEM}}}{\tau_{\text{SDEM}}} \approx 10^{-3} N_{\text{it}}$, which depend on the required number of iterations and thus indirectly on the number of particles and error tolerance. For small systems and large error tolerance, such that $N_{\text{it}} < 10^3$ is sufficient, the nonsmooth DEM will be the faster simulation method while the smooth DEM will be faster for tall columns and small error tolerance.

4. SIMULATIONS

The computational properties presented in Sec. 3 need to be verified and complemented with empirical data from numerical simulations with nonsmooth DEM. It is beyond the scope of this paper to make a thorough analysis for a range of solvers and the study is limited to projected Gauss-Seidel solver. In particular, we need to determine how the required number of iterations N_{it} for an error tolerance ϵ depend on the geometric shape and dynamic state of the system. For this purpose we perform simulations for a a number of different test systems.

We use the semi-smooth DEM as outlined in Sec. 2.4 and implemented according to the algorithm in Appendix B. The reason for this choice is that the computational complexity is identical to nonsmooth DEM and that it supports nonsmooth impact and frictional stick slip phenomena as well as elastic contact properties, in accord with Hertz law. The resulting forces and flows can therefore be compared with smooth DEM at the same time as the result in terms of computational efficiency are representative for nonsmooth DEM.

DOI: 10.1002/nme

4.1. Metrics

Specific metrics are chosen for analysing the quality of different DEM for granular matter. On a microscopic level we consider individual contacts. For any quantity, say f, we denote the mean value by $\langle \epsilon_f \rangle$ and the corresponding standard deviation by σ_f. In particular, we study the normalized contact *overlap* ϵ_{c}, *slide error* ϵ_{sl} and *slide direction error* ϵ_{dir}. For a contact α between two bodies a and b with diameter d these are defined

$$\epsilon_{\text{c}}^{(\alpha)} = g^{(\alpha)}/d \tag{40}$$

and if the contact is in slide mode $\epsilon_{\text{sl}}^{(\alpha)} = (|\vec{\mathbf{f}}_{\text{t}}^{(\alpha)}| - \mu|\vec{\mathbf{f}}_{\text{n}}^{(\alpha)}|)/\mu|\vec{\mathbf{f}}_{\text{n}}^{(\alpha)}|$ and $\epsilon_{\text{dir}} = (\theta_{\text{dir}} - \pi)/\pi$, where $\theta_{\text{dir}}^{(\alpha)} = \arccos(\vec{\mathbf{u}}_{\text{t}}^{(\alpha)}, \vec{\mathbf{f}}_{\text{t}}^{(\alpha)})$ is the angle between the directions of sliding and friction force. In the nonsmooth and semi-smooth DEM the forces on particle a are $\vec{\mathbf{f}}_{\text{n}}^{(\alpha)} = \mathbf{G}_{\text{n,[a]}}^{(\alpha)\text{T}}(:, 1:3)\lambda_{\text{n}}^{(\alpha)}/h$ and $\vec{\mathbf{f}}_{\text{t}}^{(\alpha)} = \mathbf{G}_{\text{t,[a]}}^{(\alpha)}(:, 1:3)^{\text{T}}\boldsymbol{\lambda}_{\text{t}}^{(\alpha)}/h$. We use the velocity threshold $u_{\text{thr}} \equiv 0.01\sqrt{gd/2}$ to separate between stick and slide mode $|\vec{\mathbf{u}}_{\text{t}}^{(\alpha)}| \geq u_{\text{thr}}$. We also discard the slide errors for the weakest contacts where $|\vec{\mathbf{f}}_{\text{n}}^{\alpha}| \leq f_{\text{thr}} \equiv 0.01mg$

The contact forces form *force networks*. These are weighted graphs with the particles as nodes and contact forces as edges. We use the normal force magnitude for the edge weight. The topology and force distributions of the networks are analysed.

Macroscopic fields of distribution of mass, stresses and strains are computed using *coarse graining* (or *homogenization*) [8]. In particular, we compute the mass density field $\rho(\vec{\mathbf{x}})$, velocity vector field $\vec{\mathbf{v}}(\vec{\mathbf{x}})$, strain rate tensor field $\dot{\boldsymbol{\gamma}}(\vec{\mathbf{x}})$ with norm $\dot{\gamma} = \sqrt{\text{tr}(\dot{\boldsymbol{\gamma}}^{\text{T}}\dot{\boldsymbol{\gamma}})}$, stress tensor field $\boldsymbol{\sigma}(\vec{\mathbf{x}})$ and pressure field $p(\vec{\mathbf{x}}) = \frac{1}{3}\text{tr}(\boldsymbol{\sigma})$. From these we compute the *inertial number field* $I(\vec{\mathbf{x}}) \equiv \dot{\gamma}(\vec{\mathbf{x}})d/\sqrt{p(\vec{\mathbf{x}})/\rho(\vec{\mathbf{x}})}$. The inertial number is a measure of whether a granular system is in quasistatic resting regime ($0 \leq I \ll 1$), dense flow regime ($I \lesssim 1$) or gaseous regime ($I \gtrsim 1$).

4.2. Test systems

Two class of test systems are considered: *cylindrical containers* of different size with resting granular material and *rotating drums* of different speed with dense material flow. These systems, depicted in Fig. 2, are chosen as they represent the different dynamic regimes for which both smooth and nonsmooth DEM are applicable and because they are common both in the scientific literature and in real world applications. We use spherical particles with diameter $d = 0.01$ m, mass $m = 10^{-3}$ kg corresponding to a mass density of 2500 kg/m^3, Young's modulus E $= 5 \cdot 10^6$ Pa. The equivalent spring coefficient is $k_{\text{n}} = 0.5 \cdot 10^6$ N/m. We use time step $h_{\text{NDEM}} = 10^{-2}$ s, which is 500 times larger than what is required for smooth DEM $h_{\text{SDEM}} = 5 \cdot 10^{-5}$ s. Gravity is set to $g = 9.8$ m/s^2. With these parameters the particles are practically rigid in the performed tests and any particle overlaps larger than $0.01\ d$ will be due to errors and not due to material elasticity. The friction coefficient is set to $\mu = 0.7$ and we use zero restitution coefficient $e = 0$ for both particle-particle contacts and for particle-surface contacts. Simulations are run with fixed number of iterations ranging as N_{it} ranging between 10 to 500.

4.2.1. Cylinder container Simulation is performed with particles in cylindrical container with diameter $\Phi = 1, 3, 6, 9, 15$d. The cylinders are filled with different number of particles N_{p} ranging between 5 and 100 in the $\Phi = 1d$ case, 100 to 30K in the $\Phi = 3 - 15d$ cases. The cylindrical geometry is modeled by 20 rectangular faces. The states are initialized as follows. The particles are placed in a regular cubic grid with slight perturbation. The particles are left to relax with gravity acceleration g gradually increasing from 0.01 to 9.82 m/s^2 over 60 s, sufficient to reach a stationary state. The simulation is then run for 5s. Measurements of position, velocity and contact data are made. Post-analysis is performed to obtain the metrics in Sec. 4.1. For cylindrical containers we also compute the pressure on the walls as function of height for verification of the Janssen effect. The wall pressure $p(z)$ as function of height z is computed as the average contact pressure on a cylindrical strip of width $\Delta z = 5d$ of the container wall. This procedure is repeated for 440

DOI: 10.1002/nme

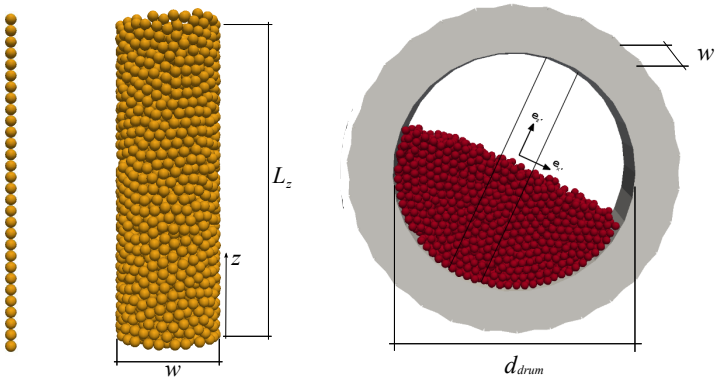

Figure 2. The test systems are cylindrical container and rotating drum.

combinations of $N_\mathrm{p}, N_\mathrm{it}$ and Φ. We treat the case $\Phi = 1d$ somewhat special. The particles are initialized in a perfect 1D column and the contacts with walls are deactivated.

4.2.2. Rotating drum For the rotating drum tests the geometric shape is kept fixed with drum diameter $\Phi = 80d$ and length $8d$. Simulations are run with drum rotation speeds $\omega = 0$, 0.03, 0.06, 0.13, 0.3, 0.63, 2.5 rad/s corresponding to dimensionless Froude number Fr = 0, 0.005, 0.013, 0.025, 0.13, 0.51, with the definition $\mathrm{Fr} = \omega\sqrt{\Phi/2g}$. The rotation is about the symmetry axis which is orthogonal to the direction of gravity. Simulations are run with different number of particles N_p ranging from 100 to 7.5K. To make the flow less dependent on the wall friction force, the drum surface is given a structural shape that of a sawtooth with 64 teeth of height $1.2d$, length $4d$ and tooth attack angle of 50 degrees. For each combination of ω, N_p and N_it the state is initialized into a stationary flow by running a 2s simulation at $N_\mathrm{it} = 500$. When changing N_it this is followed by a simulation lasting $1/3$ of an evolution plus 3.5 s. Measurements are then made during 10 consecutive time steps.

5. RESULTS AND DISCUSSION

In this section we present results of the simulation tests outlined in Sec. 4. From this we deduce a formula for the required number of iterations for projected Gauss-Seidel solver for nonsmooth (and semismooth) DEM depending on a desired error tolerance and system geometric shape and dynamics state. Supplementary data of mean penetrations, slide and friction direction, number of contacts, force networks, stress fields and velocity fields depending on number of particles and different number of iterations is provided on the web page http://umit.cs.umu.se/granular/dem/.

5.1. Cylinder container

Sample force networks from the cylindrical container test with diameter $\Phi = 9d$ is shown in Fig. 3 for $N_\mathrm{it} = 10, 100, 500$ and $N_\mathrm{p} = 7.5$K. Two main observations can be made. Firstly, too few iterations cause artificial compression of the material. For $N_\mathrm{it} = 10$ the column collapses to a height less than 40 % of the $N_\mathrm{it} = 500$ solution.

Secondly, with too few iterations the strong force chains that are a characteristic feature of granular materials do not appear. Instead, the force distributes as the hydrostatic pressure in a fluid, i.e., increases linearly with depth from the top surface. When increasing the number of iterations, strong force chain structures emerge and with this the pressure force saturate and become independent of depth in the column. This is the well-known *Janssen effect* of granular materials,

DOI: 10.1002/nme

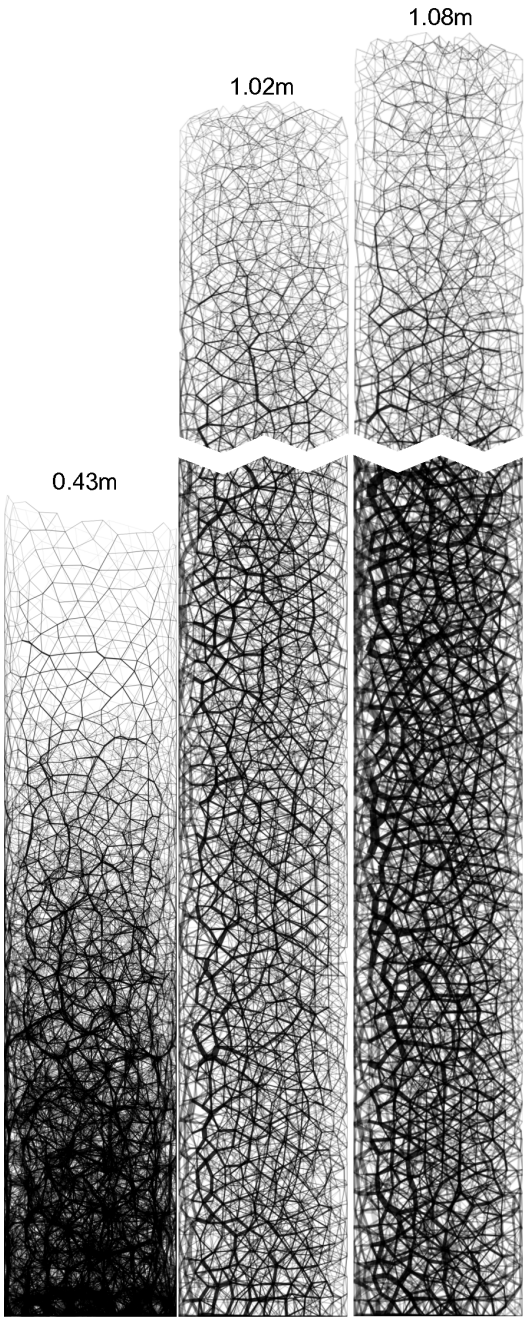

Figure 3. Sample force networks in $9d$ columns for $N_{it} = 10$, $N_{it} = 100$, $N_{it} = 500$.

which is due to an arching effect of the force chains whereby the container walls carry part of the weight of the material [20]. The Janssen force profile along the center axis for different N_{it} is shown in Fig. 4 for the $9d$ container with 7.5K particles. The pressure is normalized by the pressure $p_0 = 990$ Pa at zero height with $N_{\rm it} = 500$. It suffices with $N_{\rm it} = 50$ to capture the force saturation effect but at least $N_{\rm it} = 100$ for the correct force level of saturation.

The required number of iterations for keeping the mean penetration $\langle \epsilon_{\rm c} \rangle$ error below a given threshold is estimated from simulations and denoted by $N_{\rm it}^{\epsilon}$. The result for the $\Phi = 1d$ container is shown in Fig. 5 for different values of $\langle \epsilon_{\rm c} \rangle$ and for the cylinder size $\Phi = 3, 6, 9, 15d$ in Fig. 6,

DOI: 10.1002/nme

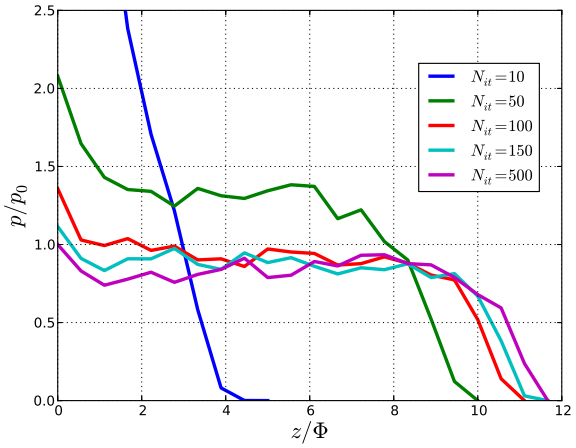

Figure 4. The Janssen force profile in a $9d$ column with 7.5K particles and different number of iterations.

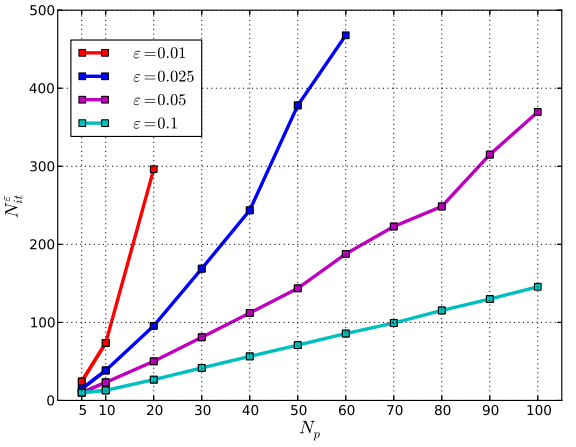

Figure 5. The required number of iterations depending on error and number of particles in a $1d$ column.

for $\langle \epsilon_c \rangle \leq 0.05$ only. In the $1d$ case the required number of iterations grow roughly linearly with the number of particles and we estimate the rate of convergence to $N_{\text{it}}^{\epsilon_1}/N_{\text{it}}^{\epsilon_2} = \epsilon_2/\epsilon_1$. In the $\Phi = 3, 6, 9, 15d$ cases in Fig. 6 we observe a *saturation effect* in the required number of iterations. Once the number of particles in the container reach above a certain number, the required number of iterations stop to increase. We discuss this effect further in Sec. 5.3. For the $3d$ container with $N_{\text{p}} \geq 1$K, we failed to create stable initial states using the described procedure.

5.2. Rotating drum

Also for the rotating drum the material behave more as a compressive fluid than a granular material with too few iterations. Sample contact force networks are displayed in Fig. 7 for $N_{\text{p}} = 7.5$K and $\omega = 0.63$ rad/s. The velocity field in the drum cross-section is presented in Fig. 8 and the flow profile $v_{x'}(z')$ in Fig. 9 in the coordinates indicated in Fig. 8. With too few iteration, $N_{\text{it}} \leq 25$, the solutions show significant artificial compression ($> 10\%$ decrease in height) and the velocity profile deviates significantly from the ones from high iterations. For notational clarity we denote the radius by $R/(\Phi/2)$. For high iterations the flow has the expected two phases: a *plug flow* zone [2], which is

DOI: 10.1002/nme

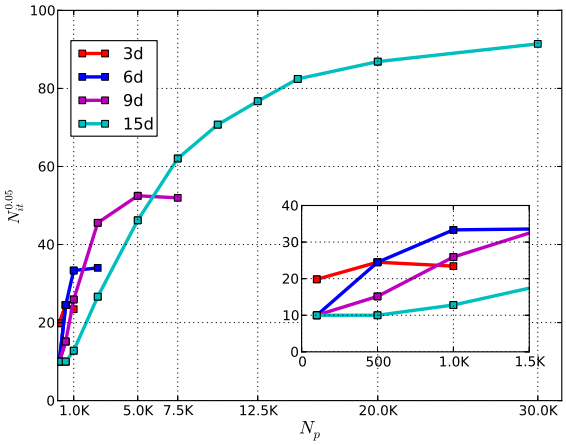

Figure 6. The required number of iterations for $\langle \epsilon_c \rangle \leq 0.05$ in the $3, 6, 9, 15d$ containers.

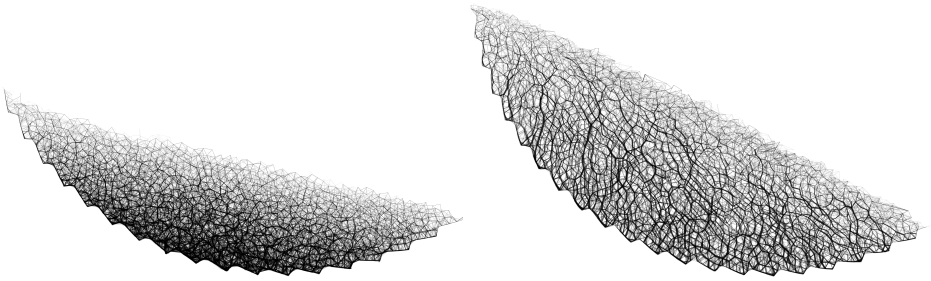

Figure 7. Force network in drum rotating with $\omega = 0.63$ rad/s, $N_p = 7.5$K for $N_{it} = 10$ (left) and 500 (right).

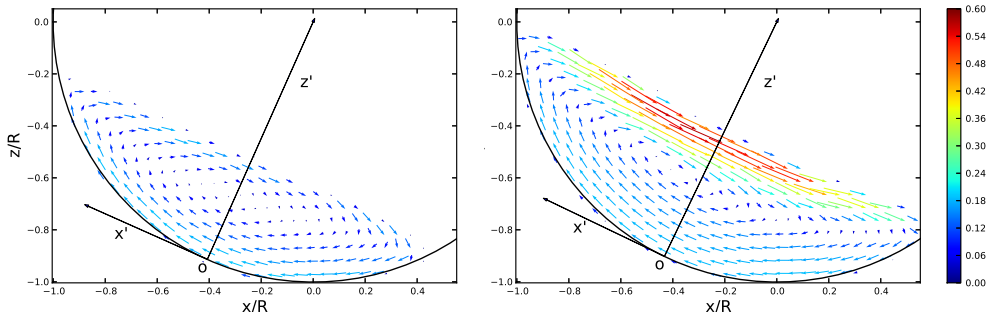

Figure 8. The velocity field in the drum rotating with $\omega = 0.63$ rad/s for $N_{it} = 10$ (left) and 500 (right).

a thick layer of material co-rotating rigidly with the drum wall, and a *shear flow* zone. The thickness of the plug flow layer is around 0.3 for $N_{\text{it}} \geq 150$ and then decrease continuously with decreasing number of iterations until all material is in shear flow. The variations in the velocity profile at high iterations, $N_{\text{it}} \geq 50$, is presumably because the flow is not entirely stationary but has a pulsation due to series of avalanches on the surface. For reference, the flow profile for rigid co-rotation with the drum is included as well as well as a solution computed with time step $h = 10^{-4}$ s and $N_{\text{it}} = 500$.

The required number of iterations $N_{\text{it}}^{0.05}$ for keeping the mean penetration error below the threshold $\langle \epsilon_c \rangle \leq 0.05$ is shown in Fig. 10 for drum speeds $\omega = 0, 0.13, 0.34, 0.63$ rad/s and different

DOI: 10.1002/nme

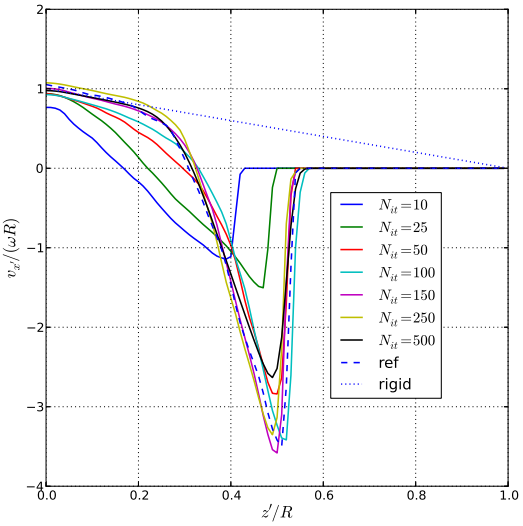

Figure 9. Velocity flow profile in the radial coordinate z' in a drum rotating with $\omega = 0.63$ rad/s and $N_{\mathrm{p}} = 7.5$K particles for different N_{it}. Rigid rotation and $h = 10^{-4}$ s reference solution is included.

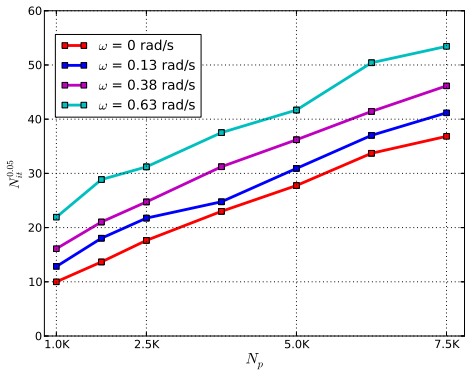

Figure 10. Number of iterations for drum with rotation speed ω and different number of particles.

number of particles. The required number of iterations scale linearly with the number of particles. Increasing drum velocity increases the required number of iterations. The fast drum $\omega = 2.5$ rad/s could not be made to converge below the error threshold $\langle \epsilon_{\mathrm{c}} \rangle \leq 0.05$. The total number of contacts is nearly independent of drum rotation speed but are redistributed from continuous contacts to impacting contacts as drum rotation speed increases. For the case of 7.5K particles and the given time-step $h_{\mathrm{NDEM}} = 0.01$, the fraction of impacting contacts is almost 50% for $\omega = 2.5$ rad/s and 10% for $\omega = 0.13$ rad/s. For $\omega = 2.5$ rad/s the impact overlaps are too large to be reduced by any number of iterations.

5.3. Computational scaling as function of geometry and state

As is clear from the results in Sec. 5.1 and 5.2 the required number of GS iterations does not depend on the number of particles or contacts alone, but on the geometric shape and size and of the dynamic state. We conjecture that the shape and size can be characterized by the system

DOI: 10.1002/nme

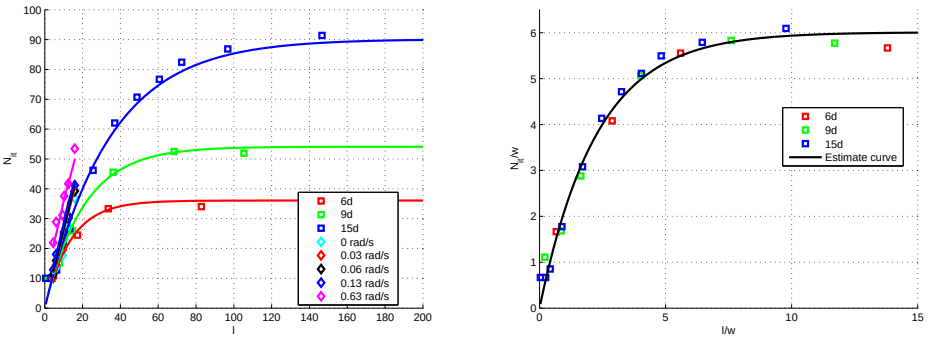

Figure 11. Number of iterations as function of system height l (left) and of shape factor l/w for containers only (right). The fitted formula is indicated as a solid line.

length l in the direction of gravity (or more dominant external force) and on the width w of the cross section area. We represent these in units of particle size d. Furthermore, we use the inertial number I for characterizing the dynamic state of the system ranging from static to rapid flow. The cylindrical container results in Fig. 6 suggest that the number of Gauss-Seidel iterations $N_{\text{it}}^{\epsilon}(l, w, I)$ scale linearly with l until it saturates at a certain length to width ratio l/w to a value proportional to width w. Supposedly, also this saturation is due to the emergence of strong force chains forming arching structure over between the container walls and thereby reducing the distance over which contact force need to be communicated by the iterative solver. In the Janssen effect, the wall pressure saturates to a value corresponding to supporting an *apparent mass* according (pp. 270 [20]) to the formula $m_{\text{app}} = m_{\text{sat}}(1 - \exp[-m_{\text{fill}}/m_{\text{sat}}])$, with actual fill mass $m_{\text{fill}} = \rho l A$, cross-section area $A = \pi w^2$ and saturation mass $m_{\text{sat}} = \rho A w / 2\mu_{\text{w}} K$ with particle-wall friction coefficient μ_{w} and Janssen coefficient $0 < K < 1$. We note that $m_{\text{app}} = m_{\text{sat}}(1 - \exp[-2\mu_{\text{w}} K l/w])$ and assume functional dependency on l/w holds for the saturation of $N_{\text{it}}^{\epsilon}(l, w, I)$ unless for systems too narrow, $w \lesssim 5$, for arching to occur. We furthermore assume that the convergence rate, $N_{\text{it}}^{\epsilon_1}/N_{\text{it}}^{\epsilon_2} = \epsilon_2/\epsilon_1$, as found for $1d$ container is a general result. The rotating drum results in Fig. 10 suggest linear dependency on rate of change. Based on these observations and assumptions, we make the following ansatz for a general system

$$N_{\text{it}}^{\epsilon}(l, w, I) = \frac{c_0(1+c_1 I)}{\epsilon} \begin{cases} w(1 - \exp[-\frac{c_2 l}{w}]) & \text{, if } w \gtrsim 5 \\ c_2 l & \text{, if } w < 5 \end{cases} \tag{41}$$

with coefficients c_0, c_1 and c_2 to be determined. Observe that in the limit where $w \gg l$ this approximates to $N_{\text{it}}^{\langle\epsilon\rangle}(l, w, I) \approx \frac{c_0(1+c_1 I)}{\epsilon} c_2 l$. In the container tests we simply identify l by the material height and w by the cylinder diameter, i.e., $w = \Phi/d$. In the drum tests we estimate l by the averaged thickness of the material in radial direction and width by w by the drum diameter. Using nonlinear least square regression for fitting Eq. (41) with the results in Fig. 6 and 10 we find the parameter values $c_0 = 0.3(0.005)$, $c_1 = 2.0(1.1)$ and $c_2 = 0.44(0.02)$ with standard error in the parenthesis. The residual of the fitting is $r^2 = 1.25$. The matching of the function compared to the test systems is displayed in Fig. 11.

With knowledge of $N_{\text{it}}^{\epsilon}(l, w, I)$ it is now possible to compare the computational efficiency of smooth and nonsmooth DEM without running any simulations but instead evaluating Eq. (36), (37), (38), (39) and Eq. (41) with $K_{NDEM}^{\text{GS}} = 10^{-6}$ from our implementation and $K_{\text{SDEM}} = 10^{-6}$ measured from running LIGGGHTS on the same computer (Appendix C). Observe that these scale factors are implementation and hardware dependent. To illustrate the computational differences we calculate the computational effort for 33 different cases. The results are found in Table I. For convenience, a calculator for the formula has been made available on web: `http://umit.cs.umu.se/granular/dem/`. In Table I we observe the smooth DEM to be considerably faster in the examples of row 13, 15, 18, 25, 27, 29 and 30 which are characterized by

DOI: 10.1002/nme

softer material, higher inertial number and tall systems. The nonsmooth DEM is considerably faster in the examples of row 1, 4-10 and 20-24 which are characterized by stiffer material, low inertial number and shallow systems. Larger error tolerance make the nonsmooth DEM more efficient. Realtime performance, i.e., when the computational effort τ/τ_{real} is smaller or equal than one, can be achieved with nonsmooth DEM in the examples of row 5, 14, 20, 26, 31 and 31 and for smooth DEM in examples 26 and 31 and with systems up to roughly 1000 contacting rigid bodies.

Table I. Computational time of nonsmooth and smooth DEM for 32 different systems. The quantities are: nomsooth DEM computational time τ_{N} and step size h_{N}, smooth DEM computational τ_{S} and step size h_{S}, number of GS iterations N_{it}, error tolerance ϵ, system width w, system height l, number of particles N_{p}, inertial number I, characteristic normal velocity v_{n} Youngs modulus E and particle diameter d.

	$\tau_{\text{N}}/\tau_{\text{S}}$	$\tau_{\text{N}}/\tau_{\text{real}}$	$\tau_{\text{S}}/\tau_{\text{real}}$	h_{N}/ms	h_{S}/ms	N_{it}	ϵ	w	l	N_{p}	I	v_{n}/mm/s	E/GPa	d/mm
1	0.02	2	100	3	0.003	19	0.05	6	10	203	0	0	0.01	1
2	0.02	10	625	3	0.003	23	0.05	15	10	1.8K	0	0	0.01	1
3	1.0	6e+3	6e+3	1	0.003	511	0.01	15	100	18K	0.1	6.1	0.01	1
4	0.02	2e+4	7e+5	3	0.003	27	0.05	500	10	2M	0.01	0.34	0.01	1
5	0.01	1	112	1	9e-5	132	0.01	1	10	10	0	0	10	1
6	0.002	5	3e+3	1	9e-5	19	0.05	6	10	280	0.01	0.29	10	1
7	0.01	200	2e+4	1	9e-5	117	0.01	15	10	1.8K	0.01	0.32	10	1
8	0.002	40	2e+4	1	9e-5	23	0.05	15	10	1.8K	0.01	0.32	10	1
9	0.05	9e+4	2e+6	1	9e-5	540	0.01	15	1000	180K	0.1	6.3	10	1
10	0.01	3e+5	2e+7	1	9e-5	131	0.01	500	10	2M	0	0	10	1
11	2.6	6e+9	2e+9	0.4	9e-5	10534	0.01	500	1000	200M	0.1	28	10	1
12	0.2	4e+8	2e+9	1	9e-5	2107	0.05	500	1000	200M	0.1	28	10	1
13	150	3e+3	20	5	0.05	13200	0.01	1	1000	1K	0	0	0.01	10
14	0.09	0.5	6	10	0.05	19	0.05	6	10	280	0	0	0.01	10
15	5.7	2e+3	400	5	0.05	511	0.01	15	100	18K	0.1	19	0.01	10
16	0.13	5e+3	4e+4	10	0.05	26	0.05	500	10	2M	0	0	0.01	10
17	0.16	6e+3	4e+4	10	0.05	32	0.05	500	10	2M	0.1	11	0.01	10
18	460	2+9	4e+6	1	0.05	10534	0.01	500	1000	200M	0.1	88	0.01	10
19	0.52	30	60	4	0.002	1320	0.01	1	100	100	0	0	10	10
20	0.007	1	200	4	0.002	19	0.05	6	10	280	0	0	10	10
21	0.07	100	2e+3	4	0.002	180	0.01	6	100	2.8K	0	0	10	10
22	0.04	500	1e+4	4	0.002	102	0.05	15	100	18K	0.1	19	10	10
23	0.01	1e+4	1e+6	4	0.002	27	0.05	500	10	2M	0.01	1.1	10	10
24	0.01	2e+4	1e+6	4	0.002	32	0.05	500	10	2M	0.1	11	10	10
25	15	2e+9	1e+8	1	0.002	10534	0.01	500	1000	200M	0.1	88	10	10
26	0.63	0.2	0.3	30	0.9	22	0.05	6	10	280	0.1	29	0.01	100
27	34	7e+3	200	10	0.9	540	0.01	15	1000	180K	0.1	63	0.01	100
28	3.0	600	200	30	0.9	108	0.05	15	1000	180K	0.1	63	0.01	100
29	2600	6e+8	2e+5	4	0.9	10534	0.01	500	1000	200M	0.1	280	0.01	100
30	100	2e+7	2e+5	20	0.9	2107	0.05	500	1000	200M	0.1	280	0.01	100
31	0.1	0.008	0.08	30	0.1	26	0.05	1	10	10	0	0	0.5	100
32	0.09	0.2	2	30	0.1	22	0.05	6	10	280	0.1	29	0.5	100
33	0.1	2	10	30	0.1	27	0.05	15	10	1.8K	0.1	32	0.5	100

For the highest material stiffness, E $= 10$ GPa, the regularization for semi-smooth DEM become too small for the system to remain well-posed and numerical instability appears as small vibrations. We compensate the high E by reducing the time step to make the regularization Σ factor in Eq. (28) remain large enough. Specifically, for E $= 10$ GPa we clamp the time step to $h_{\text{NDEM}} = 0.004$ s for $d = 0.01$ m and to $h_{\text{NDEM}} = 0.001$ s for $d = 0.001$ m. The relative velocity was estimated from the inertial number as $v_{\text{n}} \approx 0.2I\sqrt{p/\rho}$.

5.4. Effect of contact model and solver settings

The simulation results in Sec. 5 are based on the semi-smooth DEM introduced in Sec. 2.4 integrate using the projected Gauss-Seidel solver implemented following the algorithm described in Appendix B. The elastic Hertz contact model was verified in tests with 1D columns of particles deforming elastically under its weight with Young's modulus ranging from 10^5 to 10^{13} Pa and with $N_{\text{it}} = 500$. In a column of 20 particles the overlap g agrees with the Hertz contact law in Eq. (11) with mean error of less than 0.001 of a particle diameter. We also implemented and tested the nonsmooth DEM in Sec. 2.3 with regularization and constraint stabilization following to the SPOOK scheme [17]. No significant difference in convergence rate or stability was found in the comparison between this and the semi-smooth DEM.

We also investigated the effect of applying *solver warmstarting* on the convergence rate, i.e., setting the initial guess on the Lagrange multipliers equal to, or some fraction of, the value from the

DOI: 10.1002/nme

previous time step. This accelerated the convergence of 1D columns but had no significant effect on other systems. Also the effect of using sequential or random order of iterations was found to be of no significance.

6. CONCLUSIONS

The smooth and nonsmooth DEM are both be used for computing the motion and interaction forces in granular matter. The nonsmooth DEM formulation can be modified with constraints nonlinear in the gap function and mapped to the Hertz contact law with regularization and constraint stabilization terms from to conventional the viscoelastic parameters, e.g., the Young's modulus and Poisson ratio. This model, presented in Sec. 2.4 and referred to as the *semi-smooth DEM*, thus combines the potential speed-up of nonsmooth DEM with the ability of smooth DEM to accurately model viscoelastic contacts. Both the smooth and nonsmooth DEM follow as special cases in certain limits. Hence, the main difference between smooth and nonsmooth DEM lies in the required computational time for a given accuracy.

The time step size for nonsmooth (and semi-smooth) DEM is several order in magnitude larger than smooth DEM but each integration step is more computationally intense. It is difficult to judge a priori which of the methods is the most beneficial one for a given system and desired accuracy without actually testing the alternatives.

To remedy this we provide formulas for estimating the required computational time. For nonsmooth DEM with projected Gauss-Seidel solver it is given by Eq. (37) and (41). The formula takes as input system size and shape (l and w), dynamic state (I) that may range from static to rapid flow, material properties (ρ, E, ν, d) and error tolerance (ϵ) as arguments as well as the scale factors K_{NDEM}^{GS} and K_{SDEM} that depend on implementation and hardware. The relative computational efficiency of smooth and nonsmooth DEM become

$$\frac{\tau_{\mathrm{NDEM}}}{\tau_{\mathrm{SDEM}}} = \sqrt{\frac{\max\left(\frac{mv_{\mathrm{n}}^2}{\epsilon^2}, \frac{2mgd}{\epsilon^1}\right)}{kd^2}} \frac{K_{NDEM}^{GS}}{K_{\mathrm{SDEM}}} \frac{N_{\mathrm{c}}}{N_{\mathrm{p}}} \frac{c_0(1+c_1 I)}{\epsilon} \cdot \begin{cases} w(1-e^{-\frac{c_2 l}{w}}) & w \gtrsim 5 \\ c_2 l & w < 5 \end{cases} \quad (42)$$

The number of particles is estimated $N_{\mathrm{p}} \approx l \cdot w_x \cdot w_y$, $w = \max(w_x, w_y)$ and the number of contacts $N_{\mathrm{c}} = n_{\mathrm{p}} N_{\mathrm{p}}$, with $n_{\mathrm{p}} \lesssim 10$ is the average number of contact neighbors. The general trends are as follows. Smooth DEM is computationally more efficient for wide and tall systems, rapid flows and soft materials. The nonsmooth DEM with a projected Gauss-Seidel solver is more beneficial for shallow systems, static or slow flow, stiff materials and with increasing error tolerance. Examples are provided in Table I. Observe that also in the large-scale limit, $N_{\mathrm{p}} \to \infty$, it is not evident which method is the fastest, as opposed to the formula given in Ref. [4]. The nonsmooth DEM may be faster for large-scale systems if they are shallow enough. The saturation effect, in Fig. 6, on iterations at $l/w \gg c_2$ due to arching over the container walls make the projected Gauss-Seidel solver scale much better than previously thought and can make the nonsmooth DEM competitive also for tall systems.

It should be emphasized that the results are limited to the use of projected Gauss-Seidel solver for the nonsmooth DEM. This is the solver most often reported to be used for nonsmooth DEM. The scaling of the square root bracket in Eq. (42) is due to the time step size and is a general feature for the nonsmooth DEM making computational effiency increase with material stiffness and approaching the quasi-static limit. It is left for future investigations to provide empirical data for computational scaling and error analysis for other solvers that have more promising scalability, see Sec. 3. Also the effect of parallelization of smooth [31] and nonsmooth DEM [12] should be included in future work.

A third interesting topic for future research is the possibility of extending the semi-smooth DEM to an adaptive hybrid DEM, where different domains of he system is solved with a smooth and nonsmooth method and adjusted adaptively based on which method is most computational efficient based on current geometric and dynamic state. In the semi-smooth setting it is trivial to replace a

DOI: 10.1002/nme

contact constraint by the corresponding nonlinear spring force, as in Eq. (34). The computational effort become that of the nonsmooth DEM with $N_{\text{it}} = 1$.

6.1. Acknowledgements

This project was supported by LKAB, UMIT Research Lab and by Algoryx Simulations. The authors are grateful to Stefan Hedman for creating the initial analysis pipeline and performing experiments with LIGGGHTS.

APPENDIX

A. Impact stage MLCP

In large-scale simulations it become too inefficient to locate the exact time of each impact and solve for the entire contact network. Instead, fixed time step is used and impacts are detected *post facto* and solved simultaneously in a separate *impact stage* while preserving also the previously existing contact constraints.

In the impact stage we split the contact set into two. One set of impacting contacts $\mathcal{N}_{\text{i}}$, arising during the last time step, and one set of continuous contacts $\mathcal{N}_{\text{c}}$. The contact Jacobians are split correspondingly from $\mathbf{G}_{\text{n}}$ into $\mathbf{G}_{\text{in}}$ and $\mathbf{G}_{\text{cn}}$ such that $\mathbf{G}_{\text{in}}\mathbf{v}^- < 0$.

Impacts are instantaneous impulse transfers where the velocity changes discontinuously from $\mathbf{v}^-$ to $\mathbf{v}^+$ as

$$\mathbf{M}\mathbf{v}^+ = \mathbf{M}\mathbf{v}^- + \mathbf{G}_{\text{in}}^{\text{T}}\boldsymbol{\lambda}_{\text{in}} + \mathbf{G}_{\text{cn}}^{\text{T}}\boldsymbol{\lambda}_{\text{cn}} \tag{43}$$

due to the impact impulse $\mathbf{G}_{\text{in}}^{\text{T}}\boldsymbol{\lambda}_{\text{in}}$ and response $\mathbf{G}_{\text{cn}}^{\text{T}}\boldsymbol{\lambda}_{\text{cn}}$ to preserve the Signorin Coulomb law in the continuous contact network. The impacts impulses should satisfy the Newton impact law, $\mathbf{G}_{\text{in}}\mathbf{v}^+ = -e\mathbf{G}_{\text{in}}\mathbf{v}^-$, with restitution coefficient e between 0 and 1 that corresponds to completely inelastic and perfectly elastic collision, respectively. The propagation of the impulse through the network of continuous contacts should not create constraint violations, i.e., should satisfy $\mathbf{G}_{\text{cn}}\mathbf{v}^+ = 0$ and act only repulsive, $\boldsymbol{\lambda}_{\text{in}} \geq 0$. Adding tangential friction impulses satisfying the Coulomb law is straight forward as is the extension to semi-smooth DEM with regularization. Collecting the impulse equation (43) and the Newton impact law plus constraint preservation we see that they constitute a MLCP of the same form as in Eq. (19) with

$$\mathbf{H} = \begin{bmatrix} \mathbf{M} & -\bar{\mathbf{G}}_{\text{in}}^{\text{T}} & -\bar{\mathbf{G}}_{\text{cn}}^{\text{T}} & -\bar{\mathbf{G}}_{\text{it}}^{\text{T}} & -\bar{\mathbf{G}}_{\text{ct}}^{\text{T}} \\ \bar{\mathbf{G}}_{\text{in}} & \boldsymbol{\Sigma}_{\text{i}} & 0 & 0 & 0 \\ \bar{\mathbf{G}}_{\text{cn}} & 0 & \boldsymbol{\Sigma}_{\text{c}} & 0 & 0 \\ \bar{\mathbf{G}}_{\text{it}} & 0 & 0 & \boldsymbol{\Gamma}_{\text{i}} & 0 \\ \bar{\mathbf{G}}_{\text{ct}} & 0 & 0 & 0 & \boldsymbol{\Gamma}_{\text{c}} \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{v}^+ \\ \boldsymbol{\lambda}_{\text{in}} \\ \boldsymbol{\lambda}_{\text{cn}} \\ \boldsymbol{\lambda}_{\text{it}} \\ \boldsymbol{\lambda}_{\text{ct}} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -\mathbf{M}\mathbf{v}^- \\ e\mathbf{G}_{\text{in}}\mathbf{v}^- \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{44}$$

B. Projected Gauss-Seidel solver

Consider the system of equations for the nonsmooth DEM

$$\begin{bmatrix} \mathbf{M} & -\mathbf{G}^{\text{T}} \\ \mathbf{G} & \boldsymbol{\Sigma} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} \tag{45}$$

with friction cone conditions $\boldsymbol{\lambda} \in \mathcal{C}(\mu\boldsymbol{\lambda}_{\text{n}})$

$$\mathcal{C}(\mu\boldsymbol{\lambda}_{\text{n}}) \equiv \{\boldsymbol{\lambda} = [\boldsymbol{\lambda}_{\text{n}}, \boldsymbol{\lambda}_{\text{t}_1}, \boldsymbol{\lambda}_{\text{t}_1}] : \boldsymbol{\lambda}_{\text{n}} \geq 0,\ |\boldsymbol{\lambda}_{\text{t}}| \leq \mu|\boldsymbol{\lambda}_{\text{n}}|\}$$

Let the submatrices $\mathbf{M}$ and $\boldsymbol{\Sigma}$ be block diagonal and $\mathbf{G}$ block sparse. Split the linear system on Schur complement form such that

$$(\mathbf{G}\mathbf{M}^{-1}\mathbf{G}^{\text{T}} + \boldsymbol{\Sigma})\boldsymbol{\lambda} = \mathbf{q} - \mathbf{G}\mathbf{M}^{-1}\mathbf{p} \tag{46}$$

DOI: 10.1002/nme

$$\mathbf{v} = \mathbf{M}^{-1}\mathbf{p} + \mathbf{M}^{-1}\mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda} \tag{47}$$

Split the Schur matrix $\mathbf{S} = \mathbf{G}\mathbf{M}^{-1}\mathbf{G}^{\mathrm{T}} + \boldsymbol{\Sigma}$ as $\mathbf{S} = \mathbf{L} + \mathbf{D} + \mathbf{L}^{\mathrm{T}}$, where $\mathbf{D}$ is block diagonal and $\mathbf{L}$ is strictly lower triangular. On block form Eq. (50), with block indices α and β, Eq. (46) can be solved iteratively $k = 0, 1, 2, \ldots$ by

$$\mathbf{D}_{(\alpha\alpha)}\boldsymbol{\lambda}_{k+1}^{(\alpha)} + \sum_{\beta<\alpha}\mathbf{L}_{(\alpha\beta)}\boldsymbol{\lambda}_{k+1}^{(\beta)} + \sum_{\beta>\alpha}\mathbf{L}_{(\alpha\beta)}\boldsymbol{\lambda}_{k}^{(\beta)} = \mathbf{q}_{(\alpha)} - \mathbf{G}_{(\alpha)}\mathbf{M}^{-1}\mathbf{p}_{(\alpha)} \tag{48}$$

with $\mathbf{D}_{(\alpha\alpha)} = \sum_a \mathbf{G}_{[a]}^{(\alpha)}\mathbf{M}_{[aa]}^{-1}\mathbf{G}_{[a]}^{(\alpha)\mathrm{T}} + \boldsymbol{\Sigma}_{(\alpha\alpha)}$. Adding and subtracting $\mathbf{D}_{(\alpha\alpha)}\boldsymbol{\lambda}_k^{(\alpha)}$ to this we obtain the following update formula for $\boldsymbol{\lambda}_{k+1}^{(\alpha)}$

$$\mathbf{D}_{(\alpha\alpha)}\boldsymbol{\lambda}_{k+1}^{(\alpha)} + r_k^{(\alpha)} - \mathbf{D}_{(\alpha\alpha)}\boldsymbol{\lambda}_k^{(\alpha)} = 0 \tag{49}$$

where the solution must satisfy the complementarity condition $\boldsymbol{\lambda}_{k+1}^{(\alpha)} \in \mathcal{C}(\mu\boldsymbol{\lambda}_{\mathrm{n},k+1}^{(\alpha)})$ and the residual vector is

$$\mathbf{r}_k^{(\alpha)} = \mathbf{S}_{(\alpha\alpha)}\boldsymbol{\lambda}_k^{(\alpha)} + \mathbf{G}_{(\alpha)}\mathbf{M}^{-1}\mathbf{p}_{(\alpha)} - \mathbf{q}_{(\alpha)} = \mathbf{G}_{(\alpha)}\mathbf{v}' - \mathbf{q}_{(\alpha)} \tag{50}$$

where $\mathbf{v}' \equiv \mathbf{M}^{-1}\mathbf{p} + \mathbf{M}^{-1}\mathbf{G}_{(\alpha)}^{\mathrm{T}}\boldsymbol{\lambda}_k$. In the projected Gauss-Seidel we solve first the normal component. Then, if $\lambda_{\mathrm{n},k+1}^{(\alpha)} > 0$, we solve for the tangential components and project it onto the cone surface if it was outside

$$\boldsymbol{\lambda}_{\mathrm{t},k+1}^{(\alpha)} \leftarrow \mathrm{proj}_{\mu\lambda_{\mathrm{n},k+1}^{(\alpha)}}\left(\boldsymbol{\lambda}_{\mathrm{t},k+1}^{(\alpha)}\right) = \min\left(\frac{\mu\lambda_{\mathrm{n},k+1}^{(\alpha)}}{|\boldsymbol{\lambda}_{\mathrm{t},k+1}^{(\alpha)}|}, 1\right)\cdot\boldsymbol{\lambda}_{\mathrm{t},k+1}^{(\alpha)} \tag{51}$$

The algorithm is valid also for semi-smooth DEM by substituting the nonlinear normal constraints and Jacobians. For the projection onto the friction cone, replace $\mu\lambda_{\mathrm{n},k+1}^{(\alpha)}/|\boldsymbol{\lambda}_{\mathrm{t},k+1}^{(\alpha)}| \rightarrow \mu|\mathbf{f}_{\mathrm{n},k+1}^{(\alpha)}|/|\mathbf{f}_{\mathrm{t},k+1}^{(\alpha)}|$. with $\vec{\mathbf{f}}_{\mathrm{n}}^{(\alpha)} = \bar{\mathbf{G}}_{\mathrm{n}[a]}^{(\alpha)}(:, 1:3)^{\mathrm{T}}\lambda_{\mathrm{n}}^{(\alpha)}/h$ and $\vec{\mathbf{f}}_{\mathrm{t}}^{(\alpha)} = \mathbf{G}_{\mathrm{t}[a]}^{(\alpha)}(:, 1:3)^{\mathrm{T}}\boldsymbol{\lambda}_{\mathrm{t}}^{(\alpha)}/h$.
The algorithm for the projected Gauss-Seidel solver for the semi-smooth DEM then is:

Pseudocode for the algorithm is available at http://umit.cs.umu.se/granular/dem/.

C. Computer specification

The simulations where performed with a desktop computer with Intel(R) Core(TM) Xeon X5690, 3.46 GHz, 48 GB RAM on a Linux 64 bit system.

REFERENCES

[1] Algoryx Simulations. AGX multiphysics toolkit, September 2012.

[2] A. A. Boateng and P. V. Barr. Granular flow behaviour in the transverse plane of a partially filled rotating cylinder. *Journal of Fluid Mechanics*, 330:233–249, 1997.

[3] Folkmar A. Bornemann. *Homogenization in Time of Singularly Perturbed Mechanical Systems*, volume 1687 of *Lecture notes in mathematics, 0075-8434*. Springer, Berlin, 1998.

[4] L. Brendel, T. Unger, and D. E. Wolf. Contact dynamics for beginners. In H. Hinrichsen and D. E. Wolf, editors, *The Physics of Granular Media*, pages 325–343. Wiley-VCH, 2004.

[5] Nikolai V. Brilliantov, Frank Spahn, Jan-Martin Hertzsch, and Thorsten Pöschel. Model for collisions in granular gases. *Phys. Rev. E*, 53:5382–5392, May 1996.

[6] B Brogliato, Aa Ten Dam, L Paoli, F Génot, and M Abadie. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *Applied Mechanics Reviews*, 55(2):107–150, 2002.

DOI: 10.1002/nme

Projected GS MLCP solver

Input: State $(\mathbf{x}_i, \mathbf{v}_i)$, smooth force $\mathbf{f}_\text{s}$, contacts $\mathcal{N}_\text{c}$

Constants: Particle properties (m, d)

Output: New state $(\mathbf{x}_{i+1}, \mathbf{v}_{i+1}, \boldsymbol{\lambda}_{i+1})$

1. Initialization

$(\mathbf{x}, \mathbf{v}, \boldsymbol{\lambda}) = (\mathbf{x}_i, \mathbf{v}_i, 0)$

2. Contact constraint data

For every contact $\alpha \in \mathcal{N}_\text{c}$ compute:

$g_{(\alpha)}, \mathbf{G}_\text{n}^{(\alpha)}, \mathbf{G}_\text{t}^{(\alpha)}$ as in Eq. (21), (22)-(25)

$\bar{g}_{(\alpha)} = g_{(\alpha)}^{e_\text{H}}, \bar{\mathbf{G}}_\text{n}^{(\alpha)} = e_\text{H} g_{(\alpha)}^{e_\text{H}-1} \mathbf{G}_\text{n}^{(\alpha)}$

$\bar{\mathbf{G}}_\text{t} = \mathbf{G}_\text{t}$

$\boldsymbol{\Gamma}, \boldsymbol{\Sigma}, \boldsymbol{\Upsilon}$ by Eq. (30)

D_n^{-1}, D_t1^{-1} and D_t2^{-1}

3. Solve impacts

for $k = [0 : N_\text{it}]$ or until error small:

for all $\alpha \in \mathcal{N}_\text{c}$ solve normal then tangent

if α is impact then $c = 1 + e$ else $c = 1$

$r_{\text{n},k}^{(\alpha)} = \Sigma_{(\alpha\alpha)} \lambda_{\text{n},k}^{(\alpha)} - c \sum_a \bar{\mathbf{G}}_{\text{n},k[a]}^{(\alpha)} \mathbf{v}_{[a]}$

$\lambda_{\text{n},k+1}^{(\alpha)} = \lambda_{\text{n},k}^{(\alpha)} + D_{\text{n}(\alpha\alpha)}^{-1} r_{\text{n},k}^{(\alpha)}$

$\lambda_{\text{n},k+1}^{(\alpha)} = \max(0, \lambda_{\text{n},k+1}^{(\alpha)})$

$\Delta\lambda_\text{n}^{(\alpha)} = \lambda_{\text{n},k+1}^{(\alpha)} - \lambda_{\text{n},k}^{(\alpha)}$

for bodies a and b in contact α:

$\mathbf{v}_{[a]} = \mathbf{v}_{[a]} + M_{[aa]}^{-1} \bar{\mathbf{G}}_{\text{n}[a]}^{(\alpha)\text{T}} \Delta\lambda_\text{n}^{(\alpha)}$

$\mathbf{v}_{[b]} = \mathbf{v}_{[b]} + M_{[bb]}^{-1} \bar{\mathbf{G}}_{\text{n}[b]}^{(\alpha)\text{T}} \Delta\lambda_\text{n}^{(\alpha)}$

$r_{\text{t1}\alpha}^{k} = \Gamma_{(\alpha\alpha)} \lambda_{\text{t2},k}^{(\alpha)} - \sum_a \bar{\mathbf{G}}_{\text{t1},k[a]}^{(\alpha)} \mathbf{v}_{[a]}$

$r_{\text{t2}\alpha}^{k} = \Gamma_{(\alpha\alpha)} \lambda_{\text{t2},k}^{(\alpha)} - \sum_a \bar{\mathbf{G}}_{\text{t2},k[a]}^{(\alpha)} \mathbf{v}_{[a]}$

$\lambda_{\text{t1},k+1}^{(\alpha)} = \lambda_{\text{t1},k}^{(\alpha)} + D_{\text{t1}(\alpha\alpha)}^{-1} r_{\text{t1},k}^{(\alpha)}$

$\lambda_{\text{t2},k+1}^{(\alpha)} = \lambda_{\text{t2},k}^{(\alpha)} + D_{\text{t2}(\alpha\alpha)}^{-1} r_{\text{t2},k}^{(\alpha)}$

$\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)} \equiv [\lambda_{\text{t1},k+1}^{(\alpha)}, \lambda_{\text{t2},k+1}^{(\alpha)}]^\text{T}$

if $|\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)}| >= \mu |\lambda_{\text{n},k+1}^{(\alpha)}|$ then:

$\boldsymbol{\lambda}_{\text{t},(k+1)}^{(\alpha)} = \text{proj}_{\mu\lambda_{\text{n},k+1}^{(\alpha)}} (\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)})$

$\Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)} = \boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)} - \boldsymbol{\lambda}_{\text{t},k}^{(\alpha)}$

for bodies a and b in contact α:

$\mathbf{v}_{[a]} = \mathbf{v}_{[a]} + M_{[aa]}^{-1} \bar{\mathbf{G}}_{\text{t}[a]}^{(\alpha)\text{T}} \Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)}$

$\mathbf{v}_{[b]} = \mathbf{v}_{[b]} + M_{[bb]}^{-1} \bar{\mathbf{G}}_{\text{t}[b]}^{(\alpha)\text{T}} \Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)}$

4. Pre-solve

$\mathbf{q}_\text{n} = -(4/h)\boldsymbol{\Upsilon}\bar{g} + \boldsymbol{\Upsilon}\bar{\mathbf{G}}_\text{n}\mathbf{v}$

$\mathbf{p} = \mathbf{M}\mathbf{v} + h\mathbf{f}_\text{s}$

$\mathbf{v}' = \mathbf{M}^{-1}\mathbf{p}$

5. Continous contacts

for $k = [0 : N_\text{it}]$ or until error small:

for all $(\alpha) \in \mathcal{N}_\text{c}$ solve normal then tangent:

$r_{\text{n},k}^{(\alpha)} = -q_{\text{n},k}^{(\alpha)} + \Sigma_{(\alpha\alpha)} \lambda_{\text{n},k}^{(\alpha)} - \sum_a \bar{\mathbf{G}}_{\text{n},k[a]}^{(\alpha)} \mathbf{v}'_{[a]}$

$\lambda_{\text{n},k+1}^{(\alpha)} = \lambda_{\text{n},k}^{(\alpha)} + D_{\text{n}(\alpha\alpha)}^{-1} r_{\text{n},k}^{(\alpha)}$

$\lambda_{\text{n},k+1}^{(\alpha)} = \max(0, \lambda_{\text{n},k+1}^{(\alpha)})$

$\Delta\lambda_\text{n}^{(\alpha)} = \lambda_{\text{n},k+1}^{(\alpha)} - \lambda_{\text{n},k}^{(\alpha)}$

for bodies a and b in contact α:

$\mathbf{v}'_{[a]} = \mathbf{v}'_{[a]} + M_{[aa]}^{-1} \bar{\mathbf{G}}_{\text{n}[a]}^{(\alpha)\text{T}} \Delta\lambda_\text{n}^{(\alpha)}$

$\mathbf{v}'_{[b]} = \mathbf{v}'_{[b]} + M_{[bb]}^{-1} \bar{\mathbf{G}}_{\text{n}[b]}^{(\alpha)\text{T}} \Delta\lambda_\text{n}^{(\alpha)}$

$r_{\text{t1},k}^{(\alpha)} = \Gamma_{(\alpha\alpha)} \lambda_{\text{t1},k}^{(\alpha)} - \sum_a \bar{\mathbf{G}}_{\text{t1},k[a]}^{(\alpha)} \mathbf{v}'_{[a]}$

$r_{\text{t2},k}^{(\alpha)} = \Gamma_{(\alpha\alpha)} \lambda_{\text{t2},k}^{(\alpha)} - \sum_a \bar{\mathbf{G}}_{\text{t2},k[a]}^{(\alpha)} \mathbf{v}'_{[a]}$

$\lambda_{\text{t1},k+1}^{(\alpha)} = \lambda_{\text{t1},k}^{(\alpha)} + D_{\text{t1}(\alpha\alpha)}^{-1} r_{\text{t1},k}^{(\alpha)}$

$\lambda_{\text{t1},k+1}^{(\alpha)} = \lambda_{\text{t2},k}^{(\alpha)} + D_{\text{t2}(\alpha\alpha)}^{-1} r_{\text{t2},k}^{(\alpha)}$

$\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)} \equiv [\lambda_{\text{t1},k+1}^{(\alpha)}, \lambda_{\text{t2},k+1}^{(\alpha)}]^\text{T}$

if $|\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)}| >= \mu |\lambda_{\text{n},k+1}^{(\alpha)}|$ then:

$\lambda_{\text{t},k+1}^{(\alpha)} = \text{proj}_{\mu\lambda_{\text{n},k+1}^{(\alpha)}} (\boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)})$

$\Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)} = \boldsymbol{\lambda}_{\text{t},k+1}^{(\alpha)} - \boldsymbol{\lambda}_{\text{t},k}^{(\alpha)}$

for bodies a and b in contact α:

$\mathbf{v}'_{[a]} = \mathbf{v}'_{[a]} + M_{[aa]}^{-1} \bar{\mathbf{G}}_{\text{t}[a]}^{(\alpha)\text{T}} \Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)}$

$\mathbf{v}'_{[b]} = \mathbf{v}'_{[b]} + M_{[bb]}^{-1} \bar{\mathbf{G}}_{\text{t}[b]}^{(\alpha)\text{T}} \Delta\boldsymbol{\lambda}_\text{t}^{(\alpha)}$

6. Update position and orientation

$\mathbf{v}_{i+1} = \mathbf{v}'$

$\mathbf{x}_{i+1} = \mathbf{x}_i + h\mathbf{v}_{i+1}$

[7] P. A. Cundall and O. D. L. Strack. A discrete numerical model for granular assemblies. *Geotechnique*, 29:47–65, 1979.

[8] Isaac Goldhirsch. Stress, stress asymetry and couple stress: from discrete particles to contiuous fields. *Granular Matter*, 12(3):239–252, 2010.

[9] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Spring Series in Computational Mathematics*. Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, 2001.

[10] Ernst Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential Algebraic Problems*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, second revised

edition edition, 1996.

[11] Stefan Hedman. Smooth and non-smooth approaches to simulation of granular materal. Master's thesis, Department of Physics, Umeå University, Umeå, Sweden, September 2011.

[12] K. Iglberger and U. Rüde. Massively parallel granular flow simulations with non-spherical particles. *Computer Science - Research and Development*, 25:105–113, 2010.

[13] M. Jean. The non-smooth contact dynamics method. *Computer Methods in Applied Mechanics and Engineering*, 177:235–257, July 1999.

[14] Michael Jean. Numerical simulation of granular materials. In B. Cambou, M. Jean, and F. Radjai, editors, *Micromechanics of Granular Materials*, pages 149–307. John Wiley & Sons, 2004.

[15] K. Krabbenhoft, J. Huang, M.Vicente Silva, and A.V. Lyamin. Granular contact dynamics with particle elasticity. *Granular Matter*, 14:607–619, 2012.

[16] Claude Lacoursière. Splitting methods for dry frictional contact problems in rigid multibody systems: Preliminary performance results. In Mark Ollila, editor, *Conference Proceedings from SIGRAD2003, November 20–21, 2003, Umeå University, Umeå, Sweden*, pages 11–16. SIGRAD, November 2003.

[17] Claude Lacoursière. Regularized, stabilized, variational methods for multibodies. In Dag Fritzson Peter Bunus and Claus Führer, editors, *The 48th Scandinavian Conference on Simulation and Modeling (SIMS 2007), 30-31 October, 2007, Göteborg (Särö), Sweden*, pages 40–48. Linköping University Electronic Press, December 2007.

[18] Claude Lacoursière, Martin Servin, and Anders Backman. Fast and stable simulation of granular matter and machines. In Antonio Munjiza, editor, *Discrete element methods - Simulation of Discontinua: Theory and Applications*. DEM5 - The Fifth International Conference on Discrete Element Methods, London 25-26 Aug (2010), 2010.

[19] LIGGGHTS. Open source discrete element method particle simulation code, September 2012.

[20] Anita Mehta. *Granular Physics*. Cambridge University Press, New York, 2007.

[21] J. J. Moreau. Numerical aspects of the sweeping process. *Computer Methods in Applied Mechanics and Engineering*, 177:329–349, July 1999.

[22] Antonio Munjiza, editor. *The Combined Finite-discrete Element Method*. John Wiley & Sons, Ltd, Chichester, UK, 2004.

[23] Jong-Shi Pang and David E. Stewart. Differential variational inequalities. *Math. Program.*, 113(2):345–424, 2008.

[24] Thorsten Pöschel and Thomas Schwager. *Computational Granular Dynamics, Models and Algorithms*. Springer-Verlag, 2005.

[25] Farhang Radjai and Vincent Richefeu. Contact dynamics as a nonsmooth discrete element method. *Mechanics of Materials*, 41(6):715–728, June 2009.

[26] Mathieu Renouf, Vincent Acary, and Georges Dumont. 3d frictional contact and impact multibody dynamics. a comparison of algorithms suitable for real-time applications. In J.M. Goicolea, J. Cuadrado, and J.C. Garcìa Orden, editors, *Mutlibody Dynamics 2005, ECCOMAS Thematic Conference*, Madrid, Espagne, 2005.

[27] Mathieu Renouf and Pierre Alart. Conjugate gradient type algorithms for frictional multi-contact problems: applications to granular materials. *Computer Methods in Applied Mechanics and Engineering*, 194(18 - 20):2019 – 2041, 2005.

[28] Mathieu Renouf, Frederic Dubois, and Pierre Alart. A parallel version of the non smooth contact dynamics algorithm applied to the simulation of granular media. *Journal of Computational and Applied Mathematics*, 168(1-2):375 – 382, 2004.

[29] E. Rougier, A. Munjiza, and N. W. M. John. Numerical comparison of some explicit time integration schemes used in dem, fem/dem and molecular dynamics. *International Journal for Numerical Methods in Engineering*, 61(6):856–879, 2004.

[30] T. Unger, L. Brendel, D. Wolf, and J. Kertsz. Elastic behavior in contact dynamics of rigid particles. *Physical Review E*, 65(6):7, 2002.

[31] Jens H Walthe and Ivo F. Sbalzarini. Large-scale parallel discrete element simulations of granular flow. *Engineering Computations: Int J for Computer-Aided Engineering*, 26(6):688–697.

[32] Da Wang, Martin Servin, and Kjell-Ove Mickelsson. Outlet design optimization based on large-scale nonsmooth dem simulation. *Submitted for publication*, 2013.

[33] J. Williams and R. O'Connor. Discrete element simulation and the contact problem. *Archives of Computational Methods in Engineering*, 6:279–304, 1999. 10.1007/BF02818917.

[34] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method for Solid and Structural Mechanics, Sixth Edition*. Butterworth-Heinemann, 6 edition, September 2005.

DOI: 10.1002/nme