Experimental Validation of a Differential Variational Inequality-Based Approach for Handling Friction and Contact in Vehicle/Granular-Terrain Interaction

Daniel Melanz^{a,b}, Paramsothy Jayakumar^a, Dan Negrut^b

^a*U.S. Army Tank Automotive Research Development and Engineering Center, Warren, MI 48397-5000, USA*

^b*Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI 53706-1572, USA*

Abstract

The observation motivating this contribution was a perceived lack of expeditious deformable terrain models that can match in mobility analysis studies the level of fidelity delivered by today's vehicle models. Typically, the deformable terrain-tire interaction has been modeled using Finite Element Method (FEM), which continues to require prohibitively long analysis times owing to the complexity of soil behavior. Recent attempts to model deformable terrain have resorted to the use of the Discrete Element Method (DEM) to capture the soil's complex interaction with a wheeled vehicle. We assess herein a DEM approach that employs a complementarity condition to enforce non-penetration between colliding rigid bodies that make up the deformable terrain. To this end, we consider three standard terramechanics experiments: direct shear, pressure-sinkage, and single-wheel tests. We report on the validation of the complementarity form of contact dynamics with friction, assess the potential of the DEM-based exploration of fundamental phenomena in terramechanics, and identify numerical solution challenges associated with solving large-scale, quadratic optimization problems with conic constraints.

1. DEM for Terramechanics: Modeling and Numerical Solution Strategies Adopted

This contribution is motivated by an ongoing effort to identify predictive modeling approaches that can characterize dynamics of off-road vehicles. The salient feature of off-road maneuvers is the presence of deformable terrain. Owing to the complex soil behavior, deformable terrain continues to pose significant hurdles that limit the spectrum of scenarios that can be analyzed through computer modeling. The task undertaken is timely, given that it is difficult and expensive to evaluate a vehicle's performance during a majority of off-road maneuvers using physical experiments. Indeed, the range of scenarios that can be considered for physical testing is limited due to time and cost constraints. It is thus desirable to employ computer modeling in a virtual prototyping exercise that, when drawing on physics-based, predictive models, can improve designs, compress the release cycle, and reduce costs.

Email addresses: melanz@wisc.edu (Daniel Melanz), paramsothy.jayakumar.civ@mail.mil (Paramsothy Jayakumar), negrut@wisc.edu (Dan Negrut)

Figure 1: A HMMWV traversing a deformable terrain modeled using the discrete element method (DEM).

One off-road scenario of interest is shown in Fig. 1. Therein, the vehicle operates on sand, gravel, or rocky-type soil, which are modeled using DEM. DEM represents soil as a multitude of three-dimensional rigid bodies, called elements, where each element is defined by its size, shape, position, velocity, and orientation. By modeling soil using individual elements, DEM allows for significant soil deformation and transport, and the modification of properties such as soil packing structure and non-homogeneity. There are multiple formulations of DEM, classified based on how the contact and impact are handled when two bodies collide. This paper assesses the predictive attribute of the complementarity method, which models contact as a differential inclusion. To that end, it compares results obtained from standard terramechanics experiments to corresponding computer modeling analyses. These experiments include direct shear, pressure-sinkage, and single wheel tests.

1.1. Modeling Frictional Contact Via Differential Variational Inequalities

Consider a three dimensional (3D) system of rigid bodies which may interact through frictional contact. An absolute Cartesian coordinate system will be used to define the equations of motion for the time evolution of such a system [1]. Therefore, the generalized positions $\mathbf{q} = [\mathbf{r}_1^T, \epsilon_1^T, \dots, \mathbf{r}_{n_b}^T, \epsilon_{n_b}^T]^T$ and their time derivatives $\dot{\mathbf{q}} = [\dot{\mathbf{r}}_1^T, \dot{\epsilon}_1^T, \dots, \dot{\mathbf{r}}_{n_b}^T, \dot{\epsilon}_{n_b}^T]^T$ are used to describe the state of the system. Here, $\mathbf{r}_j$ is the absolute position of the center of mass of body j and ϵ_j is the quaternion used to represent rotation. Note that the angular velocity of body j in local coordinates, $\bar{\omega}_j$, may be used in place of the time derivative of the rotation quaternion. Then, the vector of generalized velocities $\mathbf{v} = [\dot{\mathbf{r}}_1^T, \bar{\omega}_1^T, \dots, \dot{\mathbf{r}}_{n_b}^T, \bar{\omega}_{n_b}^T]^T$ can be related to $\dot{\mathbf{q}}$ via a linear mapping given as $\dot{\mathbf{q}} = \mathbf{T}(\mathbf{q}) \mathbf{v}$ [1].

Due to the rigid body assumption and the choice of centroidal reference frames, the generalized mass matrix $\mathbf{M}$ is constant and diagonal. Further, let $\mathbf{f}(t, \mathbf{q}, \mathbf{v})$ be a set of generalized external forces which act on the bodies in the system. Finally, the second-order differential equations which govern the time evolution of the system can be written in the matrix form $\mathbf{M}\dot{\mathbf{v}} = \mathbf{f}(t, \mathbf{q}, \mathbf{v})$ [2].

The rigid body assumption implies that elements that come into contact should not penetrate each other. Such a condition is enforced here through unilateral constraints. To enforce the non-penetration constraint, a gap function, $\Phi(\mathbf{q}, t)$, must be defined for each pair of near-enough bodies. When two bodies are in contact, or $\Phi(\mathbf{q}, t) = 0$, a normal force acts on each of the two bodies at the contact point. When a pair of bodies is not in contact, or $\Phi(\mathbf{q}, t) > 0$, no normal force exists. This captures a complementarity condition, where one of two scenarios must hold. Either the gap is positive and the normal force is exactly zero, or vice-versa: the gap is zero, and the normal force is positive.

When a pair of bodies is in contact, friction forces may be introduced into the system through the Coulomb friction model [3]. The force associated with contact i can then be decomposed into the normal component, $\mathbf{F}_{i,N} = \hat{\gamma}_{i,n} \mathbf{n}_i$, and the tangential component, $\mathbf{F}_{i,T} = \hat{\gamma}_{i,u} \mathbf{u}_i + \hat{\gamma}_{i,w} \mathbf{w}_i$, where multipliers $\hat{\gamma}_{i,n} > 0$, $\hat{\gamma}_{i,u}$, and $\hat{\gamma}_{i,w}$ represent the magnitude of the force in each direction, $\mathbf{n}_i$ is the normal direction at the contact point, and $\mathbf{u}_i$ and $\mathbf{w}_i$ are two vectors in the contact plane such that $\mathbf{n}_i$, $\mathbf{u}_i$, and $\mathbf{w}_i$ are mutually orthonormal. Let the contact points in the local coordinates of each body be expressed as $\bar{\mathbf{s}}_{i,A}$ and $\bar{\mathbf{s}}_{i,B}$ respectively. The governing differential equations are obtained by combining the rigid body dynamics equations with the unilateral constraint equations. Then, the governing differential equations, which assume the form of a differential variational inequality (DVI) problem, become [4],

$$\dot{\mathbf{q}} = \mathbf{T}(\mathbf{q}) \mathbf{v} \quad (1)$$

$$\mathbf{M}(\mathbf{q}) \dot{\mathbf{v}} = \mathbf{f}(t, \mathbf{q}, \mathbf{v}) + \sum_{i=1}^{N_c} (\hat{\gamma}_{i,n} \mathbf{D}_{i,n}^T + \hat{\gamma}_{i,u} \mathbf{D}_{i,u}^T + \hat{\gamma}_{i,w} \mathbf{D}_{i,w}^T) \quad (2)$$

$$0 \leq \Phi_i(\mathbf{q}, t) \perp \hat{\gamma}_{i,n} \geq 0 \quad i = 1, 2, \dots, N_c \quad (3)$$

$$(\hat{\gamma}_{i,u}, \hat{\gamma}_{i,w}) = \arg \min_{\sqrt{\gamma_{i,u}^2 + \gamma_{i,w}^2} \leq \mu_i \hat{\gamma}_{i,n}} (\gamma_{i,u} \mathbf{v}^T \mathbf{D}_{i,u} + \gamma_{i,w} \mathbf{v}^T \mathbf{D}_{i,w}) . \quad (4)$$

The tangent space generators $\mathbf{D}_i = [\mathbf{D}_{i,n}, \mathbf{D}_{i,u}, \mathbf{D}_{i,w}] \in \mathbb{R}^{6n_b \times 3}$ are defined as

$$\mathbf{D}_i^T = [\mathbf{0}, \dots, -\mathbf{A}_{i,p}^T, \mathbf{A}_{i,p}^T \mathbf{A}_A \tilde{\mathbf{s}}_{i,A}, \mathbf{0}, \dots, \mathbf{0}, \mathbf{A}_{i,p}^T, -\mathbf{A}_{i,p}^T \mathbf{A}_B \tilde{\mathbf{s}}_{i,B}, \dots, \mathbf{0}] , \quad (5)$$

and are used to transform the contact forces from local to global frame, μ_i is the coefficient of friction for contact i , and N_c is the total number of possible contacts. Note that for contact between bodies A and B , the matrix $\mathbf{A}_{i,p} = [\mathbf{n}_i, \mathbf{u}_i, \mathbf{w}_i] \in \mathbb{R}^{3 \times 3}$ is used to represent the orientation of the contact in global

coordinates, and the matrices $\mathbf{A}_A$ and $\mathbf{A}_B$ as the rotation matrices of bodies A and B , respectively. where the skew-symmetric cross-product matrix is defined as

$$\tilde{\mathbf{s}} = \begin{bmatrix} 0 & -\bar{s}_z & \bar{s}_y \\ \bar{s}_z & 0 & -\bar{s}_x \\ -\bar{s}_y & \bar{s}_x & 0 \end{bmatrix}. \quad (6)$$

1.2. The Numerical Solution Scheme

Eq. 1 through 4 are discretized to obtain an approximation of the solution at discrete instants in time. In the following, superscript (l) denotes a variable at time instant $t^{(l)}$. For example, $\mathbf{q}^{(l)}$ and $\mathbf{v}^{(l)}$ represent the position and velocity at time $t^{(l)}$, respectively. Further, $\gamma_i = h\hat{\gamma}_i$ is the contact impulse for contact i . Then, the discretized form of the equations of motion is posed as [5]:

$$\mathbf{q}^{(l+1)} = \mathbf{q}^{(l)} + h\mathbf{T}(\mathbf{q}^{(l)})\mathbf{v}^{(l+1)} \quad (7)$$

$$\mathbf{M}(\mathbf{v}^{(l+1)} - \mathbf{v}^{(l)}) = h\mathbf{f}(t^{(l)}, \mathbf{q}^{(l)}, \mathbf{v}^{(l)}) + \sum_{i=1}^{N_c} (\gamma_{i,n}\mathbf{D}_{i,n}^T + \gamma_{i,u}\mathbf{D}_{i,u}^T + \gamma_{i,w}\mathbf{D}_{i,w}^T) \quad (8)$$

$$0 \leq \frac{1}{h}\Phi_i(\mathbf{q}^{(l)}, t) + \mathbf{D}_{i,n}^T\mathbf{v}^{(l+1)} \perp \gamma_{i,n} \geq 0, \quad i = 1, 2, \dots, N_c \quad (9)$$

$$(\gamma_{i,u}, \gamma_{i,w}) = \arg \min_{\sqrt{\gamma_{i,u}^2 + \gamma_{i,w}^2} \leq \mu_i \gamma_{i,n}} \left(\gamma_{i,u}\mathbf{v}^{(l+1)T}\mathbf{D}_{i,u} + \gamma_{i,w}\mathbf{v}^{(l+1)T}\mathbf{D}_{i,w} \right), \quad (10)$$

where $t^{(l+1)} = t^{(l)} + h$ for some integration time step h . Methods for solving the above problem in the most general case are lacking and various strategies are employed to simplify it. For example, the friction cones can be approximated by linear faceted pyramids, leading to a linear complementarity problem (LCP) which can be solved by pivoting or simplex methods [6]. However, these approaches in the class of direct methods can have exponential complexity in the worst-case [7]. Another artifact is the lack of isotropy, given that the friction cone is approximated by a pyramid. An alternative is to introduce a relaxation to the complementarity constraints, replacing $0 \leq (\frac{1}{h}\Phi_i(\mathbf{q}^{(l)}, t) + \mathbf{D}_{i,n}^T\mathbf{v}^{(l+1)}) \perp \gamma_{i,n} \geq 0$ with the following [8]:

$$0 \leq \left(\frac{1}{h}\Phi_i(\mathbf{q}^{(l)}, t) + \mathbf{D}_{i,n}^T\mathbf{v}^{(l+1)} - \mu_i \sqrt{(\mathbf{D}_{i,u}^T\mathbf{v}^{(l+1)})^2 + (\mathbf{D}_{i,w}^T\mathbf{v}^{(l+1)})^2} \right) \perp \gamma_{i,n} \geq 0. \quad (11)$$

As quantitatively discussed in [9] and qualitatively seen from Eq. 11, the modification is small when μ or the relative tangential velocity at the point of contact, $\sqrt{(\mathbf{D}_{i,u}^T\mathbf{v}^{(l+1)})^2 + (\mathbf{D}_{i,w}^T\mathbf{v}^{(l+1)})^2}$, are small. Additionally, it has been shown in [5] that the solution of the modified scheme will still approach the solution of the original problem as the step-size tends to zero. An iterative method has been developed

to solve this problem [10, 11]. It can be shown that solving the relaxed discretized equations of motion is equivalent to solving a Cone Complementarity Problem (CCP) of the form

$$\begin{aligned}
& \text{Find } \gamma_i^{(l+1)}, \text{ for } i = 1, \dots, N_c \\
& \text{such that } \Upsilon_i \ni \gamma_i^{(l+1)} \perp - \left(\mathbf{N} \gamma^{(l+1)} + \mathbf{r} \right)_i \in \Upsilon_i^\circ \\
& \text{where } \Upsilon_i = \{[x, y, z]^T \in \mathbb{R}^3 \mid \sqrt{y^2 + z^2} \leq \mu_i x\} \\
& \text{and } \Upsilon_i^\circ = \{[x, y, z]^T \in \mathbb{R}^3 \mid x \leq -\mu_i \sqrt{y^2 + z^2}\},
\end{aligned} \tag{12}$$

with matrix $\mathbf{N}$ and vector $\mathbf{r}$ defined as

$$\mathbf{N} = \mathbf{D}^T \mathbf{M}^{-1} \mathbf{D}, \tag{13}$$

$$\mathbf{r} = \mathbf{b} + \mathbf{D}^T \mathbf{M}^{-1} \mathbf{k}, \tag{14}$$

where $\mathbf{b}^T = [\mathbf{b}_1^T, \dots, \mathbf{b}_{N_c}^T]$ with $\mathbf{b}_i = [\frac{1}{h} \Phi_i, 0, 0]^T \in \mathbb{R}^3$, and $\mathbf{k} = \mathbf{M} \mathbf{v}^{(l)} + h \mathbf{f}(t^{(l)}, \mathbf{q}^{(l)}, \mathbf{v}^{(l)})$.

It can be verified by considering the Karush-Kuhn-Tucker first-order necessary conditions that solving the CCP of Eq. 12 is equivalent to solving a cone-constrained quadratic optimization (CCQO) problem [12]. For the case with only unilateral constraints (contacts), this optimization problem takes the form [12],

$$\begin{aligned}
\min_{\mathbf{q}} \mathbf{q}(\gamma) &= \frac{1}{2} \gamma^T \mathbf{N} \gamma + \mathbf{r}^T \gamma \\
&\text{subject to } \gamma_i \in \Upsilon_i \text{ for } i = 1, 2, \dots, N_c,
\end{aligned} \tag{15}$$

where γ_i is the triplet of multipliers associated with contact i and Υ_i is the friction cone of contact i . Note also that $\gamma = [\gamma_1^T, \gamma_2^T, \dots, \gamma_{N_c}^T]^T$.

1.3. Overall Solution Methodology and Software Infrastructure

In the adopted contact approach, the solution is produced at a sequence of time steps $t^{(0)} < t^{(1)} < \dots < t^{(N)} = T_{final}$. Currently, the integration step size is constant; i.e., $h = t_l - t_{l-1}$ is the same for any $1 \leq n \leq N$. Note that a set of initial conditions is provided at the beginning of the analysis; i.e., at time $t = t^{(0)}$. The solution is then advanced from $t^{(l)}$ to $t^{(l+1)}$ as follows. The optimization problem posed in Eq. 15 is solved to recover the set of Lagrange multipliers γ that are subsequently used in Eq. 8 to recover the new set of generalized velocities $\mathbf{v}^{(l+1)}$. Indeed, note that in the latter equation, provided γ is available, the velocities $\mathbf{v}^{(l+1)}$ are computed by multiplying from the left by $\mathbf{M}$. This matrix is constant and block diagonal. Once the velocity $\mathbf{v}^{(l+1)}$ is available, Eq. 7 is used to compute the new position and orientation of each part in the collection of bodies that make up the system of interest. It becomes manifest that the critical stage of

the solution methodology is solving the optimization of Eq. 15, a process performed at each time step t_l . To that end, we draw on a first order method called the Accelerated Projected Gradient Descent (APGD) [13], which is briefly described in the Appendix. More than 80% of the entire simulation time is spent in setting up the optimization problem (Eq. 15) and solving it via APGD.

This numerical solution strategy is implemented in an open source analysis tool called **Chrono** [14]. **Chrono**, which is an multibody dynamics engine released under a BSD3 license, produces at each time step the matrix $\mathbf{N}$ and vector $\mathbf{r}$, an operation that employs a collision detection task. For instance, in a granular material analysis that monitors the dynamics of 10 000 bodies, **Chrono** performs at each time step t_l a collision detection that would find approximately 40 000 collision events. Given that each contact event has a set of three Lagrange multipliers associated with it, **Chrono** moves next to solve the optimization problem of Eq. 15, which calls for finding the minimum of a cost function $\mathbf{q}(\gamma)$ that depends on approximately 120 000 variables. Owing to the complementarity-based formulation adopted herein, the analysis proceeds at a relatively large step size; i.e., $h \approx 0.001$ s. This translates in performing 1000 collision detections and solving 1000 optimization problems for each one second of dynamics analysis. To speed up the solution process, **Chrono** resorts to parallel computing in each stage of the analysis [15].

2. Experimental Validation

This section reports on results obtained in the process of validating the solution methodology outlined in Section §1. To this end, the **Chrono**-generated results were compared against experimental data from MIT [16] for three tests: direct shear, pressure-sinkage, and single wheel. Quikrete, a commercially available concrete mix, was used in all of the experiments for this paper. Quikrete is poorly graded, having an average particle radius of approximately 0.4 mm and bulk density $\rho_b = 1.39 \text{ g cm}^{-3}$ [17]. Based on this information, the DEM simulations used uniformly-sized spheres or ellipsoids with a major radius of 8 mm (approximately $20\times$ larger than the actual particle size) to reduce the number of bodies. Although the exact material density was not measured, the Quikrete particle density was estimated to be $\rho_g = 2.6 \text{ g cm}^{-3}$ using the following equation

$$\rho_b = M_g/V_t = \rho_g V_g/V_t$$

based on the volume of the shear box enclosure (V_t), the total mass of the granular material (M_g), and the total volume of the granular material (V_g).

2.1. Direct Shear Test

The setup for the direct shear test is shown in Fig. 2. The test is commonly used to measure the shear strength properties of a soil, specifically the cohesion, angle of friction, and shear modulus [18]. A soil

sample is placed in a shear box aligned under a load cell, which applies a normal force to the soil. The top of the shear box is clamped while the lower half can be moved in a controlled fashion by a specified displacement. The horizontal force required to displace the soil is measured to produce a plot of the shear stress as a function of shear displacement. The shear box used in the experiments had an enclosure that contained granular material and was approximately $60 \times 60 \times 60$ mm in size. Two normal pressures: 16.9 and 71.4 kPa, were tested for loosely-packed, dry soil. Four tests were performed at each normal load to determine an experimental average and standard deviation. The 3D DEM simulations were conducted by generating an upper and lower shear box using four box collision primitives. Due to the larger particle size in DEM, a larger shear enclosure ($120 \times 120 \times 120$ mm) was used for the simulation. Approximately 561 bodies were randomly generated in the interior of the two shear boxes and allowed to settle under the normal load applied by an upper box collision geometry. After settling, the lower shear box was translated at a constant rate of $6.6 \times 10^{-4} \text{ m s}^{-1}$ for a distance of 3 mm requiring $T_{final} = 4.55 \text{ s}$ for the shearing phase. Shear stress was calculated by measuring the contact force on the upper shear box and dividing by the contacted shearing area. In general, the shear stress vs. displacement relationship levels off when the material stops expanding or contracting, and when interparticle bonds are broken. The theoretical state at which the shear stress remains constant while the shear displacement increases may be called the critical state, steady state, or residual shear stress. In the case of this paper, the residual shear stress is used to study how the behavior of the direct shear simulations vary as a function of the DEM parameters, such as the friction coefficient μ .

2.1.1. Tuning of Solver Parameters

Before calibrating the physical parameters of the model, a study was performed to select Chrono solution parameters that yielded sufficiently accurate results. Indeed, if too large of an integration step size h is

Figure 2: Schematic of the direct shear test (left), Chrono simulation of the direct shear test in the sheared/final configuration (right).

used or an excessively lax convergence stopping parameter τ is adopted, the numerical results might not be “converged”, thus compromising the predictive character of the simulation. To determine the appropriate h and τ , several analyses were run using granular material composed of uniformly-sized spheres with a radius of 8 mm, a density of 2.6 g cm^{-3} , and a sliding friction coefficient of 0.5. The normal load applied was $\sigma = 16.9 \text{ kPa}$. To determine the necessary step size h , a tight solver tolerance, $\tau = 10^{-2} \text{ N}$, was used while running analyses at various h values. The results of this exercise are summarized in Fig. 3a. Conversely, given that a step size $h = 1 \times 10^{-3} \text{ s}$ appeared sufficiently small, a parametric sweep was carried out over τ to gauge the sensitivity of the analysis results with respect to the solver tolerance. The outcomes of this sweep are summarized in Fig. 3b. The results obtained elicit the following conclusions: the residual shear stress increases as the solver tolerance is tightened; the shape of the numerical shear stress profiles comes in line with expectations; and the results obtained are “converged” when $h = 1 \times 10^{-3} \text{ s}$ and $\tau = 5 \times 10^{-2} \text{ N}$.

Figure 3: Residual shear stress for the direct shear test with a varying solver time step and a fixed tolerance of $\tau = 1 \times 10^{-2} \text{ N}$ (left), the residual shear stress for the direct shear test with a varying solver tolerance and a fixed time step of $h = 1 \times 10^{-3} \text{ s}$ (right).

2.1.2. Calibration of Model Parameters

Since granular dynamics analysis is compute intensive, one would like to advance the numerical solution at a large step size and take a relatively small number of iterations to resolve each time step. The outcome of the solver parameter tuning stage described in the previous section was that the step size and convergence tolerance required for good results was $h = 1 \times 10^{-3} \text{ s}$ and $\tau = 5 \times 10^{-2} \text{ N}$, respectively.

Next, we used these solver settings to carry out a model parameter calibration aimed at determining the friction coefficient and understanding whether particle shape can improve the quality of the results. The results for the friction coefficient study are reported in Fig. 4a. The normal load applied was $\sigma = 16.9 \text{ kPa}$, the

granular material used in the experiment had density 2.6 g cm^{-3} and the shape was approximately spherical with radius 8 mm. With the exception of $\mu = 0.9$, the numerical solution obtained suggests that the shear stress profile raises as the sliding friction coefficient increases. Qualitatively, the shape of the numerical shear stress curves matches well to experimental data, with a sliding friction coefficient $\mu = 0.5$ giving the closest agreement.

Figure 4: Residual shear stress for the direct shear test with a varying friction coefficient and a constant shape ratio $s_r = 1.0$ (left). Residual shear stress for the direct shear test with a varying shape ratio and a constant friction coefficient $\mu = 0.5$ (right). The dotted and dashed lines correspond to the experimental residual shear stress average and standard deviation, respectively.

The second model parameter tuning study was performed to quantify the effect of particle shape on the direct shear test numerical results. The granular material was made up of uniformly-sized ellipsoids with a major radius of 8 mm and a varying shape ratio from 0.5 - 1.3. This ratio was defined as

$$r = \begin{cases} r_x = r, r_y = s_r \times r, r_z = r, & \text{if } s_r < 1.0 \\ r_x = r/s_r, r_y = r, r_z = r/s_r, & \text{otherwise} \end{cases}$$

This numerical experiment, which was carried out with $\mu = 0.5$, $h = 1 \times 10^{-3} \text{ s}$, $\tau = 5 \times 10^{-2} \text{ N}$, and $\sigma = 16.9 \text{ kPa}$, led to the results in Fig. 4b. Qualitatively, the shear stress curves are not very sensitive to the shape ratio. The shape of the simulated direct shear profiles matches quite well to the experimental data.

2.1.3. Predictive Attribute Assessment

For a normal load $\sigma = 16.9 \text{ kPa}$, the parameter selection process carried out in sections §2.1.1 and §2.1.2 suggests that the numerical results come close to experimental data when $\mu = 0.5$, $h = 1 \times 10^{-3} \text{ s}$, $\tau = 5 \times 10^{-2} \text{ N}$, and $s_r = 0.6$. The predictive attribute of Chrono is assessed next by keeping these solver

Figure 5: Shear stress vs. displacement for the direct shear test with $\sigma = 16.9$ kPa (left) and $\sigma = 71.4$ kPa (right).

and model parameters constant and modifying the experimental setup; i.e., the normal loading σ . Should Chrono be predictive, the numerical results for $\sigma = 71.4$ Pa would continue to be close to corresponding experimental data measured for this loading scenario. Indeed, the plot in Fig. 5 confirms that the numerical results and experimental data are similar for both loading scenarios. The direct shear simulations were run on an AMD Opteron 6274 2.2GHz processor using eight cores. For the normal load $\sigma = 16.9$ kPa, the simulation required 23.2 s of compute time per step ($23\,200\times$ slower than real time) and used an average of 4537 APGD iterations to resolve an average of 2112 collisions per step. For the normal load $\sigma = 71.4$ kPa, the simulation required 80.6 s of compute time per step ($80\,600\times$ slower than real time) and used an average of 8925 APGD iterations to resolve an average of 2129 collisions per step.

2.2. Pressure-Sinkage Test

The pressure-sinkage test, shown in Fig. 6, is used to measure the load bearing properties of a soil. To this end, a plate is pushed with a constant downward velocity to penetrate a soil sample. Designed to simulate loading rates similar to the ones exerted by a wheel during motion, the test uses a load cell to record the force induced by the plate. The soil bin used in the experiments was $150 \times 400 \times 160$ mm (W $\times$ L $\times$ H) in size. Two different plates were used to penetrate into the loosely-packed, dry soil. Both of the plates had a length of 160 mm and a height of 10 mm. Each plate had a different width; i.e., 30 and 50 mm. The plate penetrated the soil at 10 mm/s. Fifteen experimental tests were performed at each plate size. The Chrono tests were conducted by generating a soil bin in which 5377 bodies were randomly generated, dropped, and allowed to settle under gravity. After settling, a plate was moved down at a constant rate of 10 mm s^{-1} for a distance of 30 mm requiring $T_{final} = 3$ s for the pressing phase. The pressure due to sinkage was calculated by measuring the contact force on the plate and dividing by the length and width of the plate

geometry.

Figure 6: A photograph of the pressure-sinkage test experiment in the initial configuration (a), **Chrono** simulation of the pressure-sinkage test in the filled/initial configuration (b), **Chrono** simulation of the pressure-sinkage test in the pressed/final configuration (c). The colors represent the relative magnitude of the linear velocity of the bodies (red = fast, blue = slow).

2.2.1. Tuning of Solver Parameters

Before calibrating the model parameters, a study was performed to determine the necessary solver step size and termination criteria for the pressure-sinkage test. The study was performed with granular material composed of uniformly-sized spheres with a radius of 8 mm, a density of 2.6 g cm^{-3} , and $\mu = 0.5$. To determine the necessary time step, the solver tolerance was set to $\tau = 5 \times 10^{-2} \text{ N}$ and the step size h was varied to determine its effect on the numerical results. The outcomes of this study, carried out for a 50 mm plate width, are reported in Fig. 7a. The numerical integration step size $h = 1 \times 10^{-3} \text{ s}$, deemed appropriate for the shear test, yields good results for the pressure sinkage test as well. Finally, using a step size $h = 1 \times 10^{-3} \text{ s}$, a parametric sweep was carried out over τ to gauge the sensitivity of the analysis results with respect to the solver tolerance. The outcomes of this sweep are summarized in Fig. 7b. Similar to the analysis in §2.1.1, these results suggest that a solver tolerance $\tau = 5 \times 10^{-2} \text{ N}$ and a solver time step $h = 1 \times 10^{-3} \text{ s}$ is sufficient to characterize the physics of interest in the pressure-sinkage test.

Figure 7: Slope of the pressure vs. sinkage curve for various solver step sizes h (left). Slope of the pressure vs. sinkage curve for various solver tolerances τ (right).

2.2.2. Calibration of Model Parameters

The interest here is gauging the influence of the friction coefficient μ and of the particle shape, controlled through the coefficient s_r , on the numerical results of the pressure-sinkage test. The calibration tests were performed with granular material composed of uniformly-sized spheres with a radius of 8 mm, a density of 2.6 g cm^{-3} , and a varying sliding friction coefficient. In accordance with the conclusions reached in Section 2.2.1, this calibration was carried out using a simulation time step $h = 1 \times 10^{-3} \text{ s}$ and a solver tolerance $\tau = 5 \times 10^{-2} \text{ N}$. The plate width was 50 mm.

The results in Fig. 8a suggest that the slope of the pressure-sinkage curve obtained by Chrono is smaller than the experimental one by roughly a factor of five. In fact, varying the friction coefficient has very little effect on the slope of the pressure-sinkage curve. We posit that this discrepancy is due to the element dimension and/or the actual shape of the quikrete material grains. By changing the shape of the Chrono elements, for $s_r = 0.6$ the experimental and simulation results match very well. For $\mu = 0.5$, the sensitivity of the results with respect to the shape factor s_r is summarized in Fig. 8b, for s_r between 0.5 and 1.3.

2.2.3. Predictive Attribute Assessment

The next set of numerical experiments used the solver and model parameters identified in sections §2.2.1 and §2.2.2 for a plate of width 50 mm. Incidentally, these parameters assumed values identical to the ones in the shear-stress test. Specifically, the granular material was composed of uniformly-sized spheres with a radius of 8 mm, $s_r = 0.6$, $\mu = 0.5$, $\tau = 5 \times 10^{-2} \text{ N}$, and $h = 1 \times 10^{-3} \text{ seconds}$. Maintaining these parameters, the predictive attribute of Chrono was assessed by comparing the numerical results to experimental data when the width of the plate changed from 30 to 50 mm. The results obtained, summarized in Fig. 9, confirm that,

Figure 8: Slope of the pressure vs. sinkage curve for the pressure-sinkage test with a varying friction coefficient and a constant shape ratio $s_r = 1.0$ (left). Slope of the pressure vs. sinkage curve for a set of shape ratios s_r and constant friction coefficient $\mu = 0.5$ (right). The dotted and dashed lines correspond to the experimental residual shear stress average and standard deviation, respectively.

Figure 9: A comparison of the experimental and tuned pressure-sinkage profiles with a width of 50 mm(left) and a comparison of the experimental and predicted pressure-sinkage profiles with a width of 30 mm(right).

as expected, the pressure-sinkage profiles migrate to higher values as the plate width increases. Moreover, the predicted values are close to the experimental results in Fig. 9b. The pressure-sinkage simulations were run on an AMD Opteron 6274 2.2GHz processor using eight cores. For a plate of width 50 mm, the simulation required 27.9s of compute time per simulation step ($27\,900\times$ slower than real time) and used an average of 179 APGD iterations to resolve an average of 18 973 collisions per step. For a plate of width 50 mm, the simulation required 30.1s of compute time per simulation step ($30\,100\times$ slower than real time) and used an

average of 199 APGD iterations to resolve an average of 19 058 collisions per step.

2.3. Single Wheel Test in Chrono

A single wheel test is used to investigate a wheel's motion under controlled slip and normal loading conditions within a confined soil bin of dimensions $320 \times 800 \times 150$ mm ($W \times L \times H$). The drawbar pull, wheel torque, and sinkage were measured for a lug-less rigid wheel for several slip cases and loading scenarios. The wheel used in this study had a width of 160 mm and a radius $r_w = 130$ mm. To produce a desired constant slip, the wheel was rolled on the soil with a constant angular velocity of $\omega = 0.3 \text{ rad s}^{-1}$ and a certain fixed translational velocity v based on the slip defined as

$$v = (1.0 - \text{slip}) \omega r_w .$$

The numerical tests in Chrono were conducted by simulating a soil bin in which 10 790 bodies were randomly generated and allowed to settle under gravity. After settling, the wheel was rolled at the desired slip ratio for $T_{final} = 8$ s.

(a)

(b)

Figure 10: A photograph of the single wheel test experiment at MIT's Robotic Mobility Group [19] (left), Chrono simulation of the single wheel test (right).

Using the set of solver and model parameters selected in sections §2.1 and §2.2, a study was performed to gauge whether Chrono can reproduce the experimental results at varying values of wheel slips. These “predictive attribute” verification tests were performed with granular material composed of uniformly-sized ellipsoids with a major radius of 8 mm, a density of 2.6 g cm^{-3} , $s_r = 0.6$, $\mu = 0.5$, $h = 1 \times 10^{-3}$ seconds, $\tau = 5 \times 10^{-2}$ N. The quantitative results of this study are summarized in Figs. 11 through 13 for the drawbar pull, torque, and sinkage, respectively. It can be seen that as the slip of the wheel increases, the drawbar pull, torque, and sinkage also increase and the values obtained depend on the normal loads applied to the wheel; i.e., 80 N and 130 N. Moreover, the numerical values are close to the experimental results. The single wheel simulations were run on an AMD Opteron 6274 2.2GHz processor using eight cores. For a normal

load of 80 N, the simulation required 279 s of compute time per simulation step ($279\,000\times$ slower than real time) and used an average of 473 APGD iterations to resolve an average of 39 090 collisions per step. For a normal load of 130 N, the simulation required 314 s of compute time per simulation step ($314\,000\times$ slower than real time) and used an average of 513 APGD iterations to resolve an average of 39 379 collisions per step.

Figure 11: Drawbar pull vs. slip curves for a normal load 80 N (left), and 130 N (right).

Figure 12: Torque vs. slip curves for a normal load 80 N (left), and 130 N (right).

3. Conclusions and Future Research

This contribution summarizes a validation study carried out in order to assess the predictive attribute of a DVI-based frictional contact model that is expected to handle deformable terrain problems in which the

Figure 13: Sinkage vs. slip curves for a normal load 80 N (left), and 130 N (right).

soil is represented using a large number of discrete elements. The DVI model is based on complementarity conditions for contact and differential inclusions for handling friction forces. The modeling methodology has been implemented in an open source dynamics engine called **Chrono**, which is used for soft soil ground vehicle mobility studies. The conclusion of this study is that the predictive attribute of the modeling methodology, as exposed by its implementation in **Chrono**, is good. Indeed, a unique set of solution and model parameters were used to match experimental data in three tests: granular material shearing, pressure sinkage, and drawbar pull at various wheel slip levels. The simulation times are larger than hoped for, yet it is not clear whether a penalty-based DEM approach would have yielded a numerical solution more expeditiously. This simulation performance aspect, along with a better understanding of the sensitivity of simulation results with respect to the shape and size of the elements present opportunities for future research.

Acknowledgments

This work was made possible by funding provided through an US Army TARDEC ILIR grant. The last author was funded partially by the National Science Foundation grant GOALI-CMMI 1362583.

References

- [1] E. J. Haug, Computer-Aided Kinematics and Dynamics of Mechanical Systems Volume-I, Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
- [2] A. A. Shabana, Dynamics of Multibody Systems, 4th Edition, Cambridge University Press, Cambridge, England., 2013.

- [3] M. Anitescu, F. A. Potra, Formulating dynamic multi-rigid-body contact problems with friction as solvable linear complementarity problems, *Nonlinear Dynamics* 14 (1997) 231–247.
- [4] J.-S. Pang, D. Stewart, Differential variational inequalities, *Mathematical Programming* 113 (2) (2008) 345–424.
- [5] M. Anitescu, Optimization-based simulation of nonsmooth rigid multibody dynamics, *Mathematical Programming* 105 (1) (2006) 113–143. doi:<http://dx.doi.org/10.1007/s10107-005-0590-7>.
- [6] D. E. Stewart, Rigid-body dynamics with friction and impact, *SIAM Review* 42(1) (2000) 3–39.
- [7] D. Baraff, Fast contact force computation for nonpenetrating rigid bodies, in: *Computer Graphics (Proceedings of SIGGRAPH)*, 1994, pp. 23–34.
- [8] A. Tasora, A fast NCP solver for large rigid-body problems with contacts, in: C. Bottasso (Ed.), *Multibody Dynamics: Computational Methods and Applications*, Springer, 2008, pp. 45–55.
- [9] D. Melanz, M. Hammad, D. Negrut, A study of the effect of ccp-relaxation of the equations of motion in a differential variational inequality framework for handling frictional contact, Tech. Rep. TR-2015-10, Simulation-Based Engineering Laboratory, University of Wisconsin-Madison (2015).
URL <http://sbel.wisc.edu/documents/TR-2015-10.pdf>
- [10] M. Anitescu, A. Tasora, An iterative approach for cone complementarity problems for nonsmooth dynamics, *Computational Optimization and Applications* 47 (2) (2010) 207–235. doi:10.1007/s10589-008-9223-4.
- [11] A. Tasora, M. Anitescu, A matrix-free cone complementarity approach for solving large-scale, non-smooth, rigid body dynamics, *Computer Methods in Applied Mechanics and Engineering* 200 (5-8) (2011) 439–453. doi:10.1016/j.cma.2010.06.030.
- [12] T. Heyn, On the Modeling, Simulation, and Visualization of Many-Body Dynamics Problems with Friction and Contact, PhD thesis, Department of Mechanical Engineering, University of Wisconsin–Madison, http://sbel.wisc.edu/documents/TobyHeynThesis_PhDfinal.pdf (2013).
- [13] H. Mazhar, T. Heyn, A. Tasora, D. Negrut, Using Nesterov’s method to accelerate multibody dynamics with friction and contact, *ACM Trans. Graph.* 34 (3).
- [14] Project Chrono, Chrono: An Open Source Framework for the Physics-Based Simulation of Dynamic Systems, <http://www.projectchrono.org>, accessed: 2015-02-07.

- [15] D. Negrut, A. Tasora, H. Mazhar, T. Heyn, P. Hahn, Leveraging parallel computing in multibody dynamics, *Multibody System Dynamics* 27 (2012) 95–117, 10.1007/s11044-011-9262-y.
URL <http://dx.doi.org/10.1007/s11044-011-9262-y>
- [16] P. Jayakumar, D. Melanz, J. MacLennan, D. Gorsich, C. Senatore, K. Iagnemma, Scalability of classical terramechanics models for lightweight vehicle applications incorporating stochastic modeling and uncertainty propagation, *Journal of Terramechanics* 54 (2014) 37–57. doi:10.1016/j.jterra.2014.04.004.
URL <http://linkinghub.elsevier.com/retrieve/pii/S002248981400024X>
- [17] QUIKRETE - Technical Data.
URL <http://www.quikrete.com/OnTheJob/SpecData.asp>
- [18] M. G. Bekker, *Theory of land locomotion; the mechanics of vehicle mobility*, University of Michigan Press, Ann Arbor., 1956, 56010101 /L illus. 24 cm. Includes bibliography.
- [19] Robotic Mobility Group...MIT.
URL <http://web.mit.edu/mobility/index.html>
- [20] Y. Nesterov, A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$, in: *Soviet Mathematics Doklady*, Vol. 27(2), 1983, pp. 372–376.
- [21] A. Nemirovsky, D. B. Yudin, *Problem complexity and method efficiency in optimization.*, John Wiley & Sons, 1983.
- [22] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*, Vol. 87, Springer, 2003.
- [23] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences* 2 (1) (2009) 183–202.

Appendices

A. Solving the Cone-Constrained Quadratic Optimization

In [20], Nesterov developed a gradient-descent method with an improved convergence rate of $\mathcal{O}(1/k^2)$. In fact, the method in [20] was shown to be an ‘optimal’ first-order method for smooth problems [21] in terms of its performance among all first-order methods, up to a constant.

The following set of equations represents one iteration of the accelerated gradient descent (AGD) scheme [22]. Note that $\mathbf{y}_0 = \mathbf{x}_0 \in \mathbb{R}^n$, $\theta_0 = 1$, $q \in [0, 1]$ is a tuning parameter, and t_k is the step size for the current iteration.

$$\mathbf{x}_{k+1} = \mathbf{y}_k - t_k \nabla f(\mathbf{y}_k) \quad (16)$$

$$\theta_{k+1} \text{ solves } \theta_{k+1}^2 = (1 - \theta_{k+1}) \theta_k^2 + q \theta_{k+1} \quad (17)$$

$$\beta_{k+1} = \frac{\theta_k (1 - \theta_k)}{\theta_k^2 + \theta_{k+1}} \quad (18)$$

$$\mathbf{y}_{k+1} = \mathbf{x}_{k+1} + \beta_{k+1} (\mathbf{x}_{k+1} - \mathbf{x}_k) \quad (19)$$

Assume $f(\mathbf{x})$ is convex and Lipschitz continuous with constant L ; i.e., $\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_2 \leq L\|\mathbf{x} - \mathbf{y}\|_2, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then, the method described by Eqs. 16 through 19 converges for any $t_k \leq 1/L$. In the above, note that $q = 1$ leads to $\theta_k = 1$, $\beta_k = 0$, and $\mathbf{y}_k = \mathbf{x}_k$ for all $k \geq 0$, which reduces to the gradient descent method. In general, the parameter q can tune the performance of the method depending on the specifics of the objective function $f(\mathbf{x})$. For example, if $f(\mathbf{x})$ is also strongly convex, i.e., $\exists \mu > 0 : f(\mathbf{x}) \geq f(\mathbf{x}^*) + (\mu/2) \|\mathbf{x} - \mathbf{x}^*\|_2^2, \forall \mathbf{x} \in \mathbb{R}^n$, then the optimal value is $q = \mu/L$, which achieves a linear convergence rate. If the objective function is not strongly convex, or the strong convexity parameter μ is unknown, then it is often assumed that $q = 0$. Note that the original statement of the accelerated method in [20] had $q = 0$, so the convergence rate of $\mathcal{O}(1/k^2)$ is still valid.

The AGD scheme can be extended to constrained optimization by ensuring that Eq. 16 does not leave the feasible set. The resulting algorithm, called Accelerated Projected Gradient Descent (APGD) can be expressed by the following set of computations to be performed at each iteration $k \geq 0$. Once again, let $\mathbf{y}_0 = \mathbf{x}_0 \in \mathbb{R}^n$, and $\theta_0 = 1$.

$$\mathbf{x}_{k+1} = \Pi_C(\mathbf{y}_k - t_k \nabla f(\mathbf{y}_k)) \quad (20)$$

$$\theta_{k+1} \text{ solves } \theta_{k+1}^2 = (1 - \theta_{k+1}) \theta_k^2 \quad (21)$$

$$\beta_{k+1} = \frac{\theta_k (1 - \theta_k)}{\theta_k^2 + \theta_{k+1}} \quad (22)$$

$$\mathbf{y}_{k+1} = \mathbf{x}_{k+1} + \beta_{k+1} (\mathbf{x}_{k+1} - \mathbf{x}_k) \quad (23)$$

When $f(\mathbf{x})$ is convex and Lipschitz continuous with constant L , then the method described by Eqs. 20 through 23 converges for any $t_k \leq 1/L$. An equivalent algorithm was proved in [23] to converge with the same $\mathcal{O}(1/k^2)$ rate as the AGD method. These computations are performed until the residual, r , which is defined as

$$r = \|\mathbf{f}\|_2, \mathbf{f} = \frac{1}{g_d} (\gamma - \Pi_{\mathcal{K}}(\gamma - g_d(\mathbf{N}\gamma + \mathbf{r}))) \in \mathbb{R}^{3N_c}, \quad (24)$$

drops below a specified tolerance. This termination criteria is a scaled version of the projected gradient, and is designed to ensure that the computed projected gradient direction is tangent to the constraint manifold at the current iterate. This is desirable because it is known that at the optimal solution, the gradient is orthogonal to the constraint manifold. Therefore, it is logical to use the component of the gradient which is tangent to the constraint manifold as a measure of error. To understand this residual, first note that if $\gamma = \gamma^*$ is optimal, then $\Pi_{\mathcal{K}}(\gamma^* - g_d(\mathbf{N}\gamma^* + \mathbf{r})) = \gamma^*$, so $\mathbf{f} = \mathbf{0}$ and $r = 0$ as expected. Second, consider the case when γ is not optimal. Then, it can be verified that

$$\Pi_{\mathcal{K}}(\gamma - g_d(\mathbf{N}\gamma + \mathbf{r})) = \gamma - g_d\mathbf{f}. \quad (25)$$

In the preceding, the left hand side is equivalent to taking a step of length g_d in the negative gradient direction and projecting back to the feasible region. The right hand side says that the same point can be reached by taking a step of length g_d in the direction opposite of $\mathbf{f}$. In the limit, as $g_d \rightarrow 0$, the direction $\mathbf{f}$ approaches the plane tangent to the constraint manifold. Note that r could be used to measure convergence for any value of g_d , but a small value was used in practice for the reasons just stated.

A.1. APGD Algorithm

This section states the overall algorithm obtained when applying the APGD method to the problem in Eq. 15. As stated, the algorithm includes an adaptive step size which may both shrink and grow, an adaptive restart scheme based on the gradient, and a fall-back strategy to allow early termination [12].

ALGORITHM APGD($\mathbf{N}$, $\mathbf{r}$, τ , N_{max})

- (1) $\gamma_0 = \mathbf{0}_{n_c}$
- (2) $\hat{\gamma}_0 = \mathbf{1}_{n_c}$
- (3) $\mathbf{y}_0 = \gamma_0$
- (4) $\theta_0 = 1$
- (5) $L_k = \frac{\|\mathbf{N}(\gamma_0 - \hat{\gamma}_0)\|_2}{\|\gamma_0 - \hat{\gamma}_0\|_2}$
- (6) $t_k = \frac{1}{L_k}$
- (7) **for** $k := 0$ **to** N_{max}

- (8) $\mathbf{g} = \mathbf{N}\mathbf{y}_k + \mathbf{r}$
- (9) $\gamma_{k+1} = \Pi_{\mathcal{K}}(\mathbf{y}_k - t_k g)$
- (10) **while** $\frac{1}{2}\gamma_{k+1}^T \mathbf{N}\gamma_{k+1} + \gamma_{k+1}^T \mathbf{r} \geq \frac{1}{2}\mathbf{y}_k^T \mathbf{N}\mathbf{y}_k + \mathbf{y}_k^T \mathbf{r} + \mathbf{g}^T (\gamma_{k+1} - \mathbf{y}_k) + \frac{1}{2}L_k \|\gamma_{k+1} - \mathbf{y}_k\|_2^2$
- (11) $L_k = 2L_k$
- (12) $t_k = \frac{1}{L_k}$
- (13) $\gamma_{k+1} = \Pi_{\mathcal{K}}(\mathbf{y}_k - t_k g)$
- (14) **endwhile**
- (15) $\theta_{k+1} = \frac{-\theta_k^2 + \theta_k \sqrt{\theta_k^2 + 4}}{2}$
- (16) $\beta_{k+1} = \theta_k \frac{1 - \theta_k}{\theta_k^2 + \theta_{k+1}}$
- (17) $\mathbf{y}_{k+1} = \gamma_{k+1} + \beta_{k+1} (\gamma_{k+1} - \gamma_k)$
- (18) $r = r(\gamma_{k+1})$
- (19) **if** $r < \epsilon_{min}$
- (20) $r_{min} = r$
- (21) $\hat{\gamma} = \gamma_{k+1}$
- (22) **endif**
- (23) **if** $r < \tau$
- (24) **break**
- (25) **endif**
- (26) **if** $\mathbf{g}^T (\gamma_{k+1} - \gamma_k) > 0$
- (27) $\mathbf{y}_{k+1} = \gamma_{k+1}$
- (28) $\theta_{k+1} = 1$
- (29) **endif**
- (30) $L_k = 0.9L_k$
- (31) $t_k = \frac{1}{L_k}$
- (32) **endfor**
- (33) **return** Value at time step t_{l+1} , $\gamma^{l+1} := \hat{\gamma}$.