HAL
open science

An iterative approach for cone complementarity problems for nonsmooth dynamics

Mihai Anitescu, Alessandro Tasora

To cite this version:

Mihai Anitescu, Alessandro Tasora. An iterative approach for cone complementarity problems for nonsmooth dynamics. Computational Optimization and Applications, 2010, 47 (2), pp. 207-235. 10.1007/s10589-008-9223-4 . hal-01631636

HAL Id: hal-01631636

https://hal.science/hal-01631636v1

Submitted on 9 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mihai Anitescu · Alessandro Tasora

An iterative approach for cone complementarity problems for nonsmooth dynamics

December 4, 2008

Abstract Aiming at a fast and robust simulation of large multibody systems with contacts and friction, this work presents a novel method for solving large cone complementarity problems by means of a fixed-point iteration. The method is an extension of the Gauss-Seidel and Gauss-Jacobi method with overrelaxation for symmetric convex linear complementarity problems. The method is proved to be convergent under fairly standard assumptions and is shown by our tests to scale well up to 500,000 contact points and more than two millions of unknowns.

Keywords Iterative methods · cone complementarity problems · LCP · complementarity · contacts · multibody

1 Introduction

Mechanisms involving contacts and impacts between parts can be modeled in terms of multibody systems with unilateral constraints. The simulation of rigid contacts entails the solution of nonsmooth equations of motion: the dynamics is nonsmooth because of the discontinuous nature of noninterpenetration, collision, and adhesion constraints [31].

Mihai Anitescu
Mathematics and Computer Science Division,
Argonne National Laboratory,
9700 South Cass Avenue, Argonne, IL 60439, US
E-mail: anitescu@mcs.anl.gov

Alessandro Tasora
Università degli Studi di Parma,
Dipartimento di Ingegneria Industriale, 43100 Parma, Italy
E-mail: tasora@ied.unipr.it

Devices composed of rigid bodies interacting through frictional contacts are extensively used in many engineering solutions, either featuring a small number of unilateral contacts such as cam-followers and Geneva wheels or including thousands of contacts between a large number of parts, such as in the cases of palleting machines, vibratory feeders, size-segregation devices, CVT chains, and pebble reactors. Robust and efficient simulation software is mandatory, since the proper operation of these devices relies on the formation and loss of frictional contacts that cannot be easily studied by analytical methods. Given the presence of discontinuities, however, a straightforward application of numerical methods for ordinary differential equations is impracticable.

One of the most popular approaches to nonsmooth dynamics is the integrate-detect-restart method, which adopts traditional DAE or ODE integration on piecewise integrals (Caratheodory integrals) [14,15]. Nonetheless, this scheme, though reliable for systems with one degree of freedom, may fail when handling multiple unilateral constraints, because there is no way to guarantee an upper bound on the number of subproblems to solve in finite time intervals [41].

Another popular approach is represented by regularization strategies, which model contacts by means of many compliant spring-dashpot linkages. This approach requires little effort on the programming side and allows the adoption of normal ODE or DAE integrators [12,34,35,25]. Because of the high stiffness affecting the explicit integration, however, this method may require prohibitively small time steps to achieve alpha stability; moreover, the need to tune additional parameters on case-by-case basis is not welcome by end users.

These issues motivate the investigation of innovative numerical methods that can deal with multiple frictional contacts, even in case of thousands, if not millions, of moving parts. To that end, much attention was drawn by time-stepping approaches that produce weak solutions of the differential variational inequality (DVI) that describes the continuous time motion of rigid bodies with collision, contact, and friction. The DVI as a problem formulation was recently introduced in full generality and classified by differential index [29,28], though earlier numerical approaches based on DVI formulations do exist [21,20,19]. Recent work on time-stepping schemes has included both acceleration-force linear complementarity problem (LCP) approaches [8,30, 41] and velocity-impulse LCP-based time-stepping methods [38,37,5,7].

The introduction of inequalities in time-stepping schemes for DVI, coupled with a polyhedral approximation of the friction cone, leads to linear complementarity problems (LCP) [38], which are systems of complementary inequalities to be satisfied simultaneously [11]. These complex LCP problems must be solved at each time step in order to advance the integrator [19,38].

If the simulation entails a large number of contacts and rigid bodies as is the case of part feeders, packaging machines, and conveyor belts, the computational burden of classical LCP solvers can be significant. Indeed, a well-known class of approaches to LCP problems is based on *simplex methods*, also known as *direct* or *pivoting* methods, originating from the algorithms of Lemke and Dantzig [10]. However, these methods may exhibit an exponential

worst-case complexity [9]. Our experience shows that, in spite of deep optimizations [40], simplex methods still cannot practically handle multibody systems with more than one hundred colliding bodies.

In the three-dimensional case, the Coulomb friction at contact points without the use of a polyhedral approximation leads to a more complex nonlinear complementarity problem (NCP). The use of a polyhedral approximation make possible the use of typical LCP solvers [38,41,5]. Artificial anisotropy, however, affects friction because friction cones become faceted friction pyramids. In addition, such finite approximation of cones results in a far larger problem (insofar as the number of constraints) and has a negative impact on the performance of LCP solvers, which is already critical in general.

To circumvent the difficulties posed by increasing complexity of classical LCP solvers and the increased size and inaccuracy introduced by polyhedral approximation, we have developed a novel solution method, based on a fixed-point iteration with projection on a convex set, that can directly solve large cone complementarity problems with low computational overhead. The method is based on a time-stepping formulation that solves at every step a cone constrained optimization problem [1]. The time-stepping scheme was proven to converge in a measure differential inclusion sense, to the solution of the original continuous-time DVI. At every step, we solve the cone complementarity problem (CCP) that results from the optimality conditions of the cone constrained optimization problem. Note that the same formulation has been recently used as the basis of a quasistatic frictional contact model with local compliance [26] that also results in a CCP.

In systems with bilateral constraints only, our method reduces to a stationary Gauss-Seidel or Gauss-Jacobi method with successive over-relaxation [24]. For the original NCP formulation, Gauss-Seidel methods have been successfully used for thousands of rigid bodies in contact [22,17]. For these methods, however, no convergence theory exists, except for small friction coefficients, whereas our methods converges under certain conditions that do not include a small friction assumption. If the CCP is solved without overrelaxation, then our method shares certain features with the block coordinate descent method with convex constraints on the variables in a block [42].

Among the most promising applications of this method are dynamical analysis of large scenarios comprising thousands of colliding bodies, as in the case of the simulation of pebble bed nuclear reactors, granular flows, masonry stability analysis, robotics, and CAD/CAM/CAE simulations of complex devices (Fig.1), which to date are strongly limited by computational complexity issues, even on supercomputers.

2 Optimization-based Time-Stepping Scheme

In the following, we present our contact model, and we compare it to previous approaches. The object of study is a system of rigid bodies, described by state variables and contact and frictional constraints.

Fig. 1 Examples of multibody systems with many unilateral contacts.

2.1 System representation

At a time t, the position of the system is described by generalized coordinates $q(t) \in \mathbb{R}^m$ (which may include rotational coordinates that cannot be defined over a subspace homeomorphic to $\mathbb{R}^n$, for some n), and generalized velocities $v(t) \in \mathbb{R}^m$. In classical mechanics, $v(t)$ is continuous, and we can write $dq/dt = \Gamma(q)v$, where $\Gamma(q)$ is a matrix that connects the generalized velocities to the derivatives of the generalized coordinates.

In three dimensions, the position of a rigid body is described by the position x, y, z of the center of mass and a 3×3 orthogonal matrix $A \in \mathsf{SO}(3, \mathbb{R})$ that represents the rotation of a frame attached to the body with respect to a fixed-world frame. The A matrix could be function of three parameters $\in \mathbb{R}^3$ (Cardano angles, Euler angles, etc.), but this may cause numerical problems such as singularities in transformations. To overcome such problems, we adopt four-dimensional unitary quaternions $\eta \in \mathsf{S}^3 \subset \mathbb{H}$, though their space is not homeomorphic to $\mathbb{R}^3$.

Here we assume that A can be represented smoothly by three parameters, ϕ, θ, ζ: this parameterization is valid only locally, while for the global cumulative rotations we use quaternions. This reparameterization does not affect the dynamics [14]. Therefore a system with n bodies in three dimensions is represented by $m = 6n$ coordinates.

2.2 Nonpenetration constraints

Two rigid bodies should not penetrate, and, if they are in contact, there should be friction acting at the interface. To enforce the nonpenetration constraint, we assume that there exists a function $\Phi(q)$, which we call the gap function, that satisfies

$$\Phi(q) = \begin{cases} > 0 & \text{if the bodies are separated,} \\ = 0 & \text{if the bodies touch each other,} \\ < 0 & \text{if the bodies are interpenetrating.} \end{cases} \tag{1}$$

For such a function, the nonpenetration constraint becomes $\Phi(q) \geq 0$.

An example of such a mapping is the signed distance function [18], which is differentiable when the bodies are smooth and convex, at least up to some value of the interpenetration [2]. For most cases, even simple ones involving the relative position of two spheres, a differentiable signed distance function cannot be defined for all values of q. The fact that $\Phi(q)$ can be differentiably defined only on a neighborhood of the set $\Phi(q) \geq 0$ can be accommodated at the cost of making the analysis substantially more involved [3]. To simplify our discussion, we make the following assumption.

Differentiability of geometrical constraint data assumption:. Any contact is described by a gap function $\Phi(q)$ that is everywhere twice continuously differentiable.

2.3 Frictional constraints

In this work we describe the frictional constraints by conic constraints, which are an extension of complementarity models discussed in [5,38].

2.3.1 The Coulomb friction model

The model we represent and approximate is the Coulomb friction model. If a position q is feasible and the contact is active, that is, $\Phi(q) = 0$, then at the contact we have a normal force and a tangential force.

Let $\mathbf{n}$ be the normal at the contact, pointing toward the exterior of the body, and let $\mathbf{t}_1$ and $\mathbf{t}_2$ be the tangents at the contact. Here $\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2$ are mutually orthogonal vectors of length one in three dimensions. The vectors $\mathbf{n}$, $\mathbf{t}_1$, and $\mathbf{t}_2$ are a function of the position q. In the following we use v, $\mathbf{v}$ to refer to velocities, and the subscripts $_u$, $_v$ to refer to quantities related to the the two linearly independent tangential directions at a given contact.

The reaction force is impressed on the system by means of multipliers $\widehat{\gamma}_n \geq 0$, $\widehat{\gamma}_u$, and $\widehat{\gamma}_v$. The normal component of the force is $\mathbf{F}_N = \widehat{\gamma}_n \mathbf{n}$, and the tangential component of the force is $\mathbf{F}_T = \widehat{\gamma}_u \mathbf{t}_1 + \widehat{\gamma}_v \mathbf{t}_2$.

The Coulomb model consists of the following constraints:

$$\begin{aligned} \widehat{\gamma}_n \geq 0, \quad \Phi(q) \geq 0, \quad \Phi(q)\widehat{\gamma}_n = 0, \\ \mu\widehat{\gamma}_n \geq \sqrt{\widehat{\gamma}_u^2 + \widehat{\gamma}_v^2}, \quad ||\mathbf{v}_T|| \left(\mu\widehat{\gamma}_n - \sqrt{\widehat{\gamma}_u^2 + \widehat{\gamma}_v^2}\right) = 0, \\ \langle \mathbf{F}_T, \mathbf{v}_T \rangle = -\,||\mathbf{F}_T||\,||\mathbf{v}_T|| \end{aligned} \tag{2}$$

where $\mathbf{v}_T$ is the relative tangential velocity at contact. The effect of the friction over the dynamical system is defined by the friction coefficient $\mu \in \mathbb{R}^+$, that typically has a value between 0 and 1 for most materials. [1]

[1] Though the original Coulomb model distinguishes between static μ_s and kinetic μ_k friction coefficients, where usually the kinetic coefficient is slightly lower than its static counterpart, in this work we consider both to have the same value μ because the difference is not relevant for the discussion and suffices to say that a proper algorithm might adjust the friction coefficient adaptively during the simulation, depending on the slipping speed, to match complex nonlinearities in μ as a function of speed.

The first part of the constraint can be restated as

$$\mathbf{F} = \mathbf{F}_N + \mathbf{F}_T = \widehat{\gamma}_n \mathbf{n} + \widehat{\gamma}_u \mathbf{t}_1 + \widehat{\gamma}_v \mathbf{t}_2 \in \mathcal{K},$$

where $\mathcal{K}$ is a cone in three dimensions, whose slope is $\arctan(\mu)$.

The constraint $\langle \mathbf{F}_T, \mathbf{v}_T \rangle = -||\mathbf{F}_T||\,||\mathbf{v}_T||$ requires that the tangential force be opposite to the tangential velocity. This results in the reaction force being dissipative. In fact, an equivalent convenient way of expressing this constraint is by using the maximum dissipation principle [38,36,37]

$$(\widehat{\gamma}_u, \widehat{\gamma}_v) = \operatorname{argmin}_{\sqrt{\widehat{\gamma}_u^2 + \widehat{\gamma}_v^2} \leq \mu \widehat{\gamma}_n} \left(\widehat{\gamma}_u \mathbf{t}_1 + \widehat{\gamma}_v \mathbf{t}_2\right)^T \mathbf{v}_T.$$

These constraints are represented by mapping the vectors $\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2$ from contact coordinates to generalized coordinates [2].

For example, if we have a two-body system, then the generalized coordinates in the three-dimensional space are embedded in a twelve-dimensional space by using the coordinates $x_1, y_1, z_1, \phi_1, \theta_1, \zeta_1, x_2, y_2, z_2, \phi_2, \theta_2, \zeta_2$.

For a three-dimensional vector $\mathbf{v}$, the mapping to generalized coordinates is

$$\mathbf{v} \mapsto \begin{pmatrix} \mathbf{v} \\ r_1 \times \mathbf{v} \\ -\mathbf{v} \\ -r_2 \times \mathbf{v} \end{pmatrix},$$

where r_1 and r_2 are the relative positions of the contact point with respect to the centers of mass of the two bodies [2]. Using this mapping, we denote the generalized vector version of $\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2$ by D_n, D_u, D_v. One unfortunate side effect of generalized coordinates mapping is that, in the new coordinates, D_n, D_u, D_v cease to be mutually orthogonal.

If v is the generalized velocity, the tangential velocity satisfies the following

$$\mathbf{t}_1^T \mathbf{v}_T = v^T D_u, \quad \mathbf{t}_2^T \mathbf{v}_T = v^T D_v.$$

In generalized coordinates, the Coulomb model thus becomes

$$\begin{aligned} F_N = \widehat{\gamma}_n D_n, \quad F_T &= \widehat{\gamma}_u D_u + \widehat{\gamma}_v D_v, \\ \widehat{\gamma}_n \geq 0, \quad \Phi(q) \geq 0, \quad \widehat{\gamma}_n \Phi(q) &= 0, \\ \mu \widehat{\gamma}_n &\geq \sqrt{\widehat{\gamma}_u^2 + \widehat{\gamma}_v^2}, \\ \operatorname{argmin}_{\sqrt{\widehat{\gamma}_u^2 + \widehat{\gamma}_v^2} \leq \mu \widehat{\gamma}_n} \left(\widehat{\gamma}_u D_u + \widehat{\gamma}_v D_v\right)^T v &= (\widehat{\gamma}_u, \widehat{\gamma}_v). \end{aligned}$$

2.4 The overall dynamical model

The other dynamical data needed for the model are the mass matrix $M(q)$, the external force $f_e(t, q, v)$, and the term $f_c(q, v)$ which contains the effect of centrifugal and Coriolis forces. The mapping $f_c(q, v)$ is continuously differentiable and satisfies [6]

$$v^T f_c(q, v) = 0 \quad \forall q, v.$$

This equation implies that Coriolis and centrifugal forces do not provide any net work to the rigid multibody system. To simplify notation, we also make the following assumption.

Constant mass matrix assumption: The mass matrix $M(q) \in \mathbb{R}^{m\times m}$ is positive definite and constant. This assumption is satisfied in two dimensions and three dimensions if we use the Newton-Euler formulation in body coordinates [23].

With this definition, we can define the total force

$$f_t(t,q,v) = f_e(t,q,v) + f_c(q,v). \tag{3}$$

Assume now that we have p potential contact constraints, which are enforced by the nonpenetration constraints $\Phi^i(q) \geq 0$, $i = 1,2,\ldots,p$.

In the following, we denote by the superscript i the data associated to the potential contact i. The continuous model is the following differential variational inequality [29]:

$$\begin{aligned}
M\frac{dv}{dt} &= \sum_{i=1,2,\ldots,p} \left(\widehat{\gamma}_n^i D_n^i + \widehat{\gamma}_u^i D_u^i + \widehat{\gamma}_v^i D_v^i\right) + f_t(t,q,v) \\
\frac{dq}{dt} &= \Gamma(q)\, v \\
\widehat{\gamma}_n^i &\geq 0 \quad \perp \quad \Phi^i(q) \geq 0, \quad i = 1,2,\ldots,p \\
\left(\widehat{\gamma}_u^i, \widehat{\gamma}_v^i\right) &= \operatorname{argmin}_{\mu^i\widehat{\gamma}_n^i \geq \sqrt{(\widehat{\gamma}_u^i)^2+(\widehat{\gamma}_v^i)^2}} \left(\widehat{\gamma}_u D_u^i + \widehat{\gamma}_v D_v^i\right)^T v, \quad i = 1,2,\ldots,p.
\end{aligned} \tag{4}$$

Here $\Gamma(q)$ is used to transform the generalized velocities into derivatives of the generalized positions. For instance, when dealing with rotations, $\Gamma(q)$ can be a linear mapping from three-dimensional angular speeds into four-dimensional time derivatives of unit quaternions.

The Coulomb model used in this work is the predominant model used in the engineering literature to describe dry friction. Unfortunately, the model may be inconsistent: there exist configurations for which the model does not have a solution [8,37]. This situation has led to the need to explore weaker formulations where the forces are measures and Newton's law is satisfied in a measure differential inclusion sense [37]. It has been shown that solutions in that sense do exist and can be found by time stepping schemes [36].

We will consider all collisions that appear during the simulation of the inelastic type. Therefore, they are naturally treated by the time-stepping scheme through a change of active set without the need to modify the algebraic expression of the scheme.

2.5 Time stepping scheme

We now define a stepping scheme for the continuous time formulation. We start at the time $t^{(l)}$, position $q^{(l)}$, and velocity $v^{(l)}$ with time step h. The

scheme is expressed by the following equation problem with equilibrium constraints:

$$
\begin{aligned}
M(v^{(l+1)} - v^l) = & \sum_{j \in \mathcal{A}(q^{(l)},\epsilon)} \left(\gamma_n^i D_n^i + \gamma_u^i D_u^i + \gamma_v^i D_v^i\right) \\
& + h f_t(t^{(l)}, q^{(l)}, v^{(l)}) && (5) \\
0 \leq & \frac{1}{h}\Phi^i(q^{(l)}) + \nabla \Phi^{i^T} v^{(l+1)} && (6) \\
& \perp \quad \gamma_n^i \geq 0,\ i \in \mathcal{A}(q^{(l)},\epsilon) \\
\left(\gamma_u^i, \gamma_v^i\right) = & \operatorname{argmin}_{\mu^i \gamma_n^i \geq \sqrt{(\gamma_u^i)^2 + (\gamma_v^i)^2}} \left[v^T(\gamma_u D_u^i + \gamma_v D_v^i)\right] && (7) \\
& i \in \mathcal{A}(q^{(l)},\epsilon) \\
q^{(l+1)} = & q^{(l)} + h\Gamma(q^{(l)}) v^{(l+1)}, && (8)
\end{aligned}
$$

where

$$
\mathcal{A}(q,\epsilon) = \left\{ i \,\middle|\, i \in \{1,2,\ldots,p\},\ \ \Phi^i(q) \leq \epsilon \right\}. \tag{9}
$$

We have denoted by γ_s the constraint impulses of a contact constraint, that is, $\gamma_s = h\widehat{\gamma}_s$, for $s = n, u, v$. The $\frac{1}{h}\Phi^i(q^{(l)})$ term achieves constraint stabilization and its effect is amply discussed in [3].

In previous work, we have shown that the scheme is convergent, as the time step h goes to 0 to the solution of a measure differential inclusion [1]. Solutions of the subproblems, when the nonlinear constraint is approximated by a piecewise linear cone, can be found by Lemke's algorithm [5]. Nonetheless, in [4] we have also demonstrated that, as the number of constraints in the problem increases, the computational cost of Lemke's method increases far faster than linearly with the size of the problem. As an alternative we proposed to solve the problem as a monotone optimization by introducing a relaxation over the complementarity constraints, that is, we modified the time-stepping scheme by replacing the equation (6) with:

$$
\begin{aligned}
0 \leq & \frac{1}{h}\Phi^i(q^{(l)}) + \nabla \Phi^{i^T} v^{(l+1)} \\
& -\mu^i \sqrt{(D_u^{i,T} v)^2 + (D_v^{i,T} v)^2} \perp \quad \gamma_n^i \geq 0,\ i \in \mathcal{A}(q^{(l)},\epsilon).
\end{aligned} \tag{10}
$$

We note that the modified formulation does fit the paradigm put forth in [27]. We also note, however, that, at least in its most evident obvious formulations the problem would violate either assumption (H2) or assumption (H4) that are used in that reference to prove existence of solutions, though it is also clear that even mild regularizations would satisfy those assumptions.

In this work, we do not discuss elastic or partially inelastic collision. This is equivalent to considering a 0 restitution coefficient. It is conceivable that the approach in [32,5] can be adapted to our scheme if restitution is needed, but that is an issue that needs further study, due to the effective normal compliance that this scheme presents, as discussed below. What ever the modification, however, it is likely to result in a cone complementarity one-step problem whose matrix has similar structure to the one considered here. How to solve that problem, as opposed to the time stepping scheme itself, is the focus of this article.

We have shown in [1] that, as $h \to 0$, the solution of the modified time-stepping scheme will approach the solution of the same measure differential inclusion as the original scheme. In addition, we have shown that the iterates produced by the modified scheme approach the ones of the original scheme provided that $\mu^i \gamma_n^i \sqrt{(D_u^{i,T} v)^2 + (D_v^{i,T} v)^2} << 1$ [4]. We remark that this regime is precisely the one in which pebble bed simulators operate [33], an application that motivates our second example, as well as other granular flow applications.

The physical meaning of the modified scheme was described to some extent in [1], and is connected to the microscopic description of contact with friction as rigid piecewise linear asperities whose tangent of the side angle is equal to the macroscopic friction coefficient. The effect of the modification on the iterates of the time stepping scheme is that it allows for some amount of normal motion at the contact even when the body is supposed to be in contact. In some sense, the scheme allows for a "boundary layer", effectively normal compliance, at the contact whose size is proportional to $\mu \mathbf{v}_T h$ (but that otherwise does not depend on any parameter, as penalty schemes do). This interpretation can be demonstrated on a two-dimensional example assuming persistent contact that is already established. Of course, the modification (10) may not be suitable for all simulations with contact and friction and a good example of an unsuitable configuration is the case where two bodies are in initial contact with large tangential velocity, as is discussed in [1].

For the rest of the paper we use the modified scheme, which uses (10) instead of (6).

2.6 Cone complementarity formulation

If we now write the optimality conditions for the equilibrium constraint in (7), we obtain that, for any $i \in \mathcal{A}(q^{(l)}, \epsilon)$, there exists a Lagrange multiplier λ^i such that

$$\lambda^i \gamma_u^i = -D_u^{i,T} v,\ \lambda^i \gamma_v^i = -D_v^{i,T} v,\ \lambda^i \geq 0 \perp \mu^i \gamma_n^i - \sqrt{(\gamma_u^i)^2 + (\gamma_v^i)^2} \geq 0. \quad (11)$$

The first two equations imply that $\lambda^i \sqrt{(\gamma_u^i)^2 + (\gamma_v^i)^2} = \sqrt{\left(D_u^{i,T} v\right)^2 + \left(D_v^{i,T} v\right)^2}$, while the last equation implies that

$$0 = \lambda^i \sqrt{(\gamma_u^i)^2 + (\gamma_v^i)^2} \left(\mu^i \gamma_n^i - \sqrt{(\gamma_u^i)^2 + (\gamma_v^i)^2}\right)$$

and, in turn, that

$$\mu^i \gamma_n^i \sqrt{\left(D_u^{i,T} v\right)^2 + \left(D_v^{i,T} v\right)^2} = \lambda^i \left((\gamma_u^i)^2 + (\gamma_v^i)^2\right). \quad (12)$$

We now define, for $i \in \mathcal{A}(q^l, \epsilon)$ the vectors

$$u^i = (u_1^i, u_2^i, u_3^i), \quad w^i = (\gamma_n^i, \gamma_u^i, \gamma_v^i)$$
$$u_1^i = \frac{1}{h}\Phi^i(q^{(l)}) + \nabla\Phi^{i^T} v^{(l+1)},\ u_2^i = D_u^{i,T} v,\ u_3^i = D_v^{i,T} v.$$

We calculate the scalar product

$$\begin{aligned} u^{i^T} w^i &= \gamma_n^i \left(\frac{1}{h}\Phi^i(q^{(l)}) + \nabla\Phi^{i^T} v^{(l+1)}\right) + \gamma_u^i D_u^{i,T} v + \gamma_v^i D_v^{i,T} v \\ &\overset{(10),(11)}{=} \mu^i \gamma_n^i \sqrt{\left(D_u^{i,T} v\right)^2 + \left(D_v^{i,T} v\right)^2} - \lambda^i \left(\left(\gamma_u^i\right)^2 + \left(\gamma_v^i\right)^2\right) \\ &= 0, \end{aligned}$$

which implies that

$$u^{i^T} w^i = 0, \text{ and thus } u^i \perp w^i. \tag{13}$$

We now define the cones

$$\Lambda^i = \left\{(x, y, z) \in \mathbb{R}^3 | x \geq \mu^i \sqrt{y^2 + z^2}\right\}, \quad \mathcal{FC}^i = \left\{(x, y, z) \in \mathbb{R}^3 | \mu^i x \geq \sqrt{y^2 + z^2}\right\}$$

It immediately follows that Λ^i is the negative polar cone of $\mathcal{FC}^i$, that is, $\tilde{u} \in \Lambda^i$ and $\tilde{w} \in \mathcal{FC}^i$ imply that $\tilde{u}^T \tilde{w} \geq 0$. Then (7), (10), and (13) imply that the following set of cone complementarity constraint holds:

$$-u^i \in \mathcal{FC}^{i\circ} \perp w^i \in \mathcal{FC}^i, \quad i \in \mathcal{A}(q^{(l)}, \epsilon), \tag{14}$$

where we denote by C° the polar cone of a given cone C, that is, $C^\circ = \{x \in \mathbb{R}^m | \langle x, y\rangle \leq 0,\ \forall y \in C\}$.

We now define the vector

$$\tilde{k}^{(l)} = M v^{(l)} + h f_t(t^{(l)}, q^{(l)}, v^{(l)}). \tag{15}$$

Then, equations (15) and (14), together with (5) and the definition of the vectors u^i and w^i, result in the following problem:

$$\begin{aligned} M v^{(l+1)} &= \tilde{k}^{(l)} + \sum_{i \in \mathcal{A}(q^{(l)}, \epsilon)} \left(\gamma_n^i D_n^i + \gamma_u^i D_u^i + \gamma_v^i D_v^i\right), \\ i \in \mathcal{A}(q^{(l)}, \epsilon) \quad & \left(\tfrac{1}{h}\Phi^i(q^{(l)}) + \nabla\Phi^{i^T} v^{(l+1)}, D_u^{i,T} v^{(l+1)}, D_v^{i,T} v^{(l+1)}\right) \\ & \in -\mathcal{FC}^{i\circ} \perp (\gamma_n^i, \gamma_u^i, \gamma_v^i) \in \mathcal{FC}^i. \end{aligned} \tag{16}$$

We denote by $n_{\mathcal{A}}$ the number of elements in the set $\mathcal{A}(q^l, \epsilon)$. We then define the following vectors:

$$\begin{aligned} b \in \mathbb{R}^{3n_{\mathcal{A}}} &= \left(\tfrac{1}{h}\Phi^{i_1}(q^{(l)}), 0, 0, \tfrac{1}{h}\Phi^{i_2}(q^{(l)}), 0, 0, \ldots, \tfrac{1}{h}\Phi^{i_{n_{\mathcal{A}}}}(q^{(l)}), 0, 0\right) \\ d \in \mathbb{R}^{3n_{\mathcal{A}}} &= \Big(\tfrac{1}{h}\Phi^{i_1}(q^{(l)}) + D_n^{i_1^T} M^{-1}\tilde{k}^{(l)}, D_u^{i_1^T} M^{-1}\tilde{k}^{(l)}, D_v^{i_1^T} M^{-1}\tilde{k}^{(l)}, \\ &\quad \tfrac{1}{h}\Phi^{i_2}(q^{(l)}) + D_n^{i_2^T} M^{-1}\tilde{k}^{(l)}, D_u^{i_2^T} M^{-1}\tilde{k}^{(l)}, D_v^{i_2^T} M^{-1}\tilde{k}^{(l)}, \\ &\quad \ldots, \tfrac{1}{h}\Phi^{i_{n_{\mathcal{A}}}}(q^{(l)}) + D_n^{i_{n_{\mathcal{A}}}^T} M^{-1}\tilde{k}^{(l)}, D_u^{i_{n_{\mathcal{A}}}^T} M^{-1}\tilde{k}^{(l)}, D_v^{i_{n_{\mathcal{A}}}^T} M^{-1}\tilde{k}^{(l)}\Big) \\ \gamma \in \mathbb{R}^{3n_{\mathcal{A}}} &= \left(\gamma_n^{i_1}, \gamma_u^{i_1}, \gamma_v^{i_1}, \gamma_n^{i_2}, \gamma_u^{i_2}, \gamma_v^{i_2}, \ldots, \gamma_n^{i_{n_{\mathcal{A}}}}, \gamma_u^{i_{n_{\mathcal{A}}}}, \gamma_v^{i_{n_{\mathcal{A}}}}\right) \end{aligned} \tag{17}$$

and the following matrices

$$\begin{aligned} D^i &= \left[D^i_n, D^i_u, D^i_v\right], i \in \mathcal{A}(q^{(l)}, \epsilon), \\ D &= \left[D^{i_1}, D^{i_2}, \ldots, D^{i_{n_\mathcal{A}}}\right], \quad N = D^T M^{-1} D. \end{aligned} \tag{18}$$

Note that the matrix N is positive semidefinite.

In addition, for a vector $\tilde{d} \in \mathbb{R}^{3n_\mathcal{A}}$, we define by $\tilde{d}^i \in \mathbb{R}^3 = (\tilde{d}_{3*(i-1)+1}, \tilde{d}_{3*(i-1)+2}, \tilde{d}_{3*i})$. Note that $\tilde{d}^i$ is a vector, whereas $\tilde{d}_i$ is a real number component. This convention allows us, after multiplying with M^{-1} its first equation, to write the problem (16) as the conic complementarity problem

$$(N\gamma + d)^i \in -\mathcal{FC}^{i\circ} \perp \gamma^i \in \mathcal{FC}^i, i = 1, 2, \ldots, n_\mathcal{A}. \tag{19}$$

3 Convergence Theory of the Iterative Method

We now describe the structure of projection operators over direct sums of cones. Assume that we have a set of closed convex cones $\Upsilon^i \subset \mathbb{R}^{n_i}$, where the index takes the values $i = 1, 2, \ldots, n_k$. We consider the Cartesian product of such cones $\Upsilon = \bigoplus_{i=1}^{n_k} \Upsilon^i$, which we assume is a cone in $\mathbb{R}^{n_c}$, that is, that the sum of the dimensions of the element cones satisfies $n_c = \sum\limits_{i=1}^{n_k} n_i$. In this section and in the sequel, for a vector $x \in \mathbb{R}^{n_c}$, we denote by x_i, $i = 1, 2, \ldots, n_k$ its components that satisfy $x_i \in \mathbb{R}^{n_i}$, that is, $x = (x_1, x_2, \ldots, x_{n_k})$. Since all the operations we will carry out will be on blocks corresponding to the partition of x into its components x_i, there will be no confusion between x_i and the components of x. Note that the Cartesian product cone is also a convex cone. Note that we have chosen to use subscripts to denote indices of blocks of the vector x in order to avoid collusion with iteration indices. When particularizing the results to the case of the cone complementarity problem (19) we will again use superscripts for variables γ pertaining to a contact with index i.

For a convex cone, $C \subset \mathbb{R}^m$, we denote by $\Pi_C(y)$ the projection of the vector $y \in \mathbb{R}^m$ onto the convex cone C. From the theory of convexity, it follows that the projection has the following properties.

P1 $\|\Pi_C(y_1) - \Pi_C(y_2)\|^2 \leq \langle \Pi_C(y_1) - \Pi_C(y_2), y_1 - y_2 \rangle, \quad \forall y_1, y_2 \in \mathbb{R}^m$ [16][Proposition 3.1.3].
P2 $x = \Pi_C(y) \Leftrightarrow x \in C,\ y - x \in C^\circ,\ \langle x, y - x \rangle = 0$ [16][Proposition 3.2.3].
P3 $\Pi_\Upsilon(x) = (\Pi_{\Upsilon^1}(x_1), \Pi_{\Upsilon^2}(x_2), \ldots, \Pi_{\Upsilon^{n_k}}(x_{n_k}))$
P4 $\Upsilon^\circ = \bigoplus_{i=1}^{n_k} \Upsilon^{i,\circ}$

The last two properties are a straightforward application of the properties of convex cones and their projections.

Consider now the symmetric positive semidefinite matrix N. We define the following cone complementarity problem:

$$(CCP) \quad s_i = (Nx + q)_i \in -\Upsilon^{i,\circ},\ x_i \in \Upsilon^i,\ \langle x_i, s_i \rangle = 0,\ i = 1, 2, \ldots, n_k. \tag{20}$$

It is immediate that it represents the optimality conditions of the following optimization problem with conic constraints:

$$(OC)\begin{array}{ll}\min & f(x)=\frac{1}{2}x^TNx+q^Tx \\ s.t. & x_i\in\Upsilon^i, \qquad\qquad i=1,2,\ldots,n_k.\end{array}$$

The goal of this section is to analyze the following iterative method. We start with an arbitrarily chosen initial point $x^0\in\Upsilon$. The iterative method is defined by the formula

$$\begin{array}{l}x^{r+1}=\lambda\Pi_{\Upsilon}\left(x^r-\omega B^r\left(Nx^r+q+K^r\left(x^{r+1}-x^r\right)\right)\right)+(1-\lambda)\,x^r, \\ \qquad r=0,1,2,\ldots,\end{array} \tag{21}$$

where λ,ω are parameters that satisfy $0<\lambda\leq 1,\quad \omega>0$; for each r, the matrix K^r is a strictly block upper triangular or strictly block lower triangular, with blocks corresponding to the partition of the vector $x\in\mathbb{R}^{n_c}$ into the components x_i as outlined in the beginning of the section. In addition, B^r is a positive diagonal matrix, which is made of identity blocks whose sizes correspond to the same partition of the vector x. We therefore have

$$B^r=\begin{pmatrix}\eta_1 I_{n_1} & 0 & \cdots 0\\ 0 & \eta_2 I_{n_2} & \cdots 0\\ \vdots & \vdots & \ddots \vdots\\ 0 & 0 & \cdots \eta_{n_k} I_{n_{n_k}}\end{pmatrix},\quad L^r=\begin{pmatrix}0 & K_{12} & K_{13} & \cdots & K_{1n_k}\\ 0 & 0 & K_{23} & \cdots & K_{2n_k}\\ 0 & 0 & 0 & \cdots & K_{3n_k}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 0\end{pmatrix}, \tag{22}$$

where $\eta_i>0,\quad i=1,2,\ldots,n_k$, $I_{n_i}\in\mathbb{R}^{n_i\times n_i}$, $K_{ij}\in\mathbb{R}^{n_i\times n_j}$, $1\leq i<j\leq n_k$, and we have either that $K^r=L^r$, or that $K^r=L^{rT}$.

We will use the following assumptions.

A1 The matrix N of the problem (CCP) is symmetric and positive semi-definite.

A2 There exists a positive number, $\alpha>0$ such that, at any iteration r, $\quad r=0,1,2,\ldots$, we have that $B^r\succ\alpha I$

A3 There exists a positive number, $\beta>0$ such that, at any iteration r, $\quad r=0,1,2,\ldots$, we have that $(x^{r+1}-x^r)^T\left((\lambda\omega B^r)^{-1}+K^r-\frac{N}{2}\right)(x^{r+1}-x^r)\geq\beta\left\|x^{r+1}-x^r\right\|^2$.

To analyze the convergence behavior of the iteration (21), we used the same approach as Murty [24], adapted to the case of general convex cones. We first characterize the solution of the cone complementarity problem in terms of a fixed point of an appropriate mapping.

Theorem 1 *Assume that B is a positive definite diagonal matrix with a block structure prescribed in (22). The vector $x\in\Upsilon$ is a solution of the cone complementarity problem (CCP) if and only if it satisfies the following fixed-point relationship:*

$$\Pi_{\Upsilon}\left(x-\omega B\left(Nx+q\right)\right)=x.$$

Proof By the second property of projections [P2] we have that the vector x satisfies the fixed-point relationship if and only if it satisfies the following relationships:

$$x \in \Upsilon, \quad (x - \omega B\,(Nx + q)) - x = s \in \Upsilon^{\circ}, \quad \langle x, s\rangle = 0.$$

In turn, these are equivalent to

$$x_i \in \Upsilon, \quad (-\omega B\,(Nx + q))_i = s_i \in \Upsilon^{\circ}, \quad \langle x_i, s_i\rangle = 0, \quad i = 1, 2, \ldots, n_k.$$

From the property (P3) of the cones and the fact that the diagonal matrix B has the structure described in (22), such that the blocks corresponding to the components of x are multiples of the identity, it immediately follows that

$$(-\omega B\,(Nx + q))_i = \eta_i \omega\,(-\,(Nx + q)_i,)\,,$$

where we have used the notation from (22). In turn, this implies that the previously displayed equation is equivalent to

$$x_i \in \Upsilon, \quad (-\,(Nx + q))_i = s_i \in \Upsilon^{\circ}, \quad \langle x_i, s_i\rangle = 0, \quad i = 1, 2, \ldots, n_k,$$

which is precisely (CCP). The proof is complete. □

Theorem 2 *Assume that B is a positive definite matrix with the structure described in (22). Then $\forall x \in \mathbb{R}^{n_c}$ we have that*

$$(\Pi_{\Upsilon}(x) - x)^T B^{-1}\,(\Pi_{\Upsilon}(x) - y) = \left\langle (\Pi_{\Upsilon}(x) - x)\,, B^{-1}\,(\Pi_{\Upsilon}(x) - y)\right\rangle \leq 0, \quad \forall y \in \Upsilon.$$

Proof From the definition of the total cone Υ we have that $\Upsilon = \bigoplus_{i=1}^{n_k} \Upsilon^i$. Since the matrix B is diagonal with the structure described in (22), we immediately have that

$$\begin{aligned}\left\langle (\Pi_{\Upsilon}(x) - x)\,, B^{-1}\,(\Pi_{\Upsilon}(x) - y)\right\rangle &= \sum_{i=1}^{n_k} \frac{1}{\eta_i} \left\langle (\Pi_{\Upsilon}(x) - x)_i\,, (\Pi_{\Upsilon}(x) - y)_i\right\rangle \\ &= \sum_{i=1}^{n_k} \frac{1}{\eta_i} \left\langle \Pi_{\Upsilon^i}(x_i) - x_i, \Pi_{\Upsilon^i}(x_i) - y_i\right\rangle.\end{aligned}$$

The last relation follows from the property [P3] of the cones and projections onto them. It is therefore sufficient to show that

$$x_i \in \mathbb{R}^{n_i} \Rightarrow \langle \Pi_{\Upsilon^i}(x_i) - x_i, \Pi_{\Upsilon^i}(x_i) - y_i\rangle \leq 0, \quad \forall y_i \in \Upsilon^i. \tag{23}$$

Using property [P1] of the cones, we have that

$$\langle \Pi_{\Upsilon^i}(x_i) - y_i, \Pi_{\Upsilon^i}(x_i) - y_i\rangle \leq \langle \Pi_{\Upsilon^i}(x_i) - y_i, x_i - y_i\rangle\,, \quad \forall y_i \in \Upsilon^i, x_i \in \mathbb{R}^{n_i}.$$

Using the fact that the scalar product is a bilinear form and taking the term from the right to the left with a change sign, we obtain that

$$\langle \Pi_{\Upsilon^i}(x_i) - y_i, \Pi_{\Upsilon^i}(x_i) - x_i\rangle \leq 0, \quad \forall y_i \in \Upsilon^i, x_i \in \mathbb{R}^{n_i},$$

which proves the equation (23) and therefore the theorem. The proof is complete. □

Theorem 3 *Let $\{x^r : r = 1, 2, \ldots\}$ be the sequence of points obtained under the iterative scheme (21). Assume that $x^0 \in \Upsilon$ and that the sequences of matrices B^r and K^r are bounded. Then we have that*

$$f(x^{r+1}) - f(x^r) \leq -\beta \left\| x^{r+1} - x^r \right\|^2$$

for any iteration index r, and any accumulation point of the sequence x^r is a solution of (CCP).

Proof The proof is identical to the proof from Murty [24], where the projection on the positive orthant, $+$, is replaced with the projection on the cone Π_Υ. The only property of the projection that is used is the one from Theorem 2, which holds for the general case as well. We nonetheless include it here for completeness.

Since the initial point satisfies $x^0 \in \Upsilon$ and from (21), we conclude that $x^r \in \Upsilon$, $\forall r = 1, 2, \ldots$ From straightforward manipulation it follows that

$$f\left(x^{r+1}\right) - f\left(x^r\right) = \omega B^r \left(Nx^r + q\right)^T \left(\omega B^r\right)^{-1} \left(x^{r+1} - x^r\right)$$
$$+ \left(x^{r+1} - x^r\right) N \frac{\left(x^{r+1} - x^r\right)}{2} \left(\frac{\left(x^{r+1} - (1-\lambda)x^r\right)}{\lambda} - x^r + \omega B^r \left(Nx^r + q + K^r \left(x^{r+1} \text{ -}\right.\right.\right.$$
$$\left(\omega B^r\right)^{-1} \left(x^{r+1} - x^r\right) + \left(x^{r+1} - x^r\right) \left(\tfrac{N}{2} - \left(\lambda\omega B^r\right)^{-1} - K^r\right) \left(x^{r+1} - x^r\right) =$$
$$\lambda \left(\frac{\left(x^{r+1} - (1-\lambda)x^r\right)}{\lambda} - \left(x^r - \omega B^r \left(Nx^r + q + K^r \left(x^{r+1} - x^r\right)\right)\right) \right)^T$$
$$\left(\omega B^r\right)^{-1} \left(\frac{\left(x^{r+1} - (1-\lambda)x^r\right)}{\lambda} - x^r \right) + \left(x^{r+1} - x^r\right) \left(\tfrac{N}{2} - \left(\lambda\omega B^r\right)^{-1} - K^r\right) \left(x^{r+1} \text{ -}\right.$$

From (21) we know that

$$\frac{\left(x^{r+1} - (1-\lambda)\, x^r\right)}{\lambda} = \Pi_\Upsilon \left(x^r - \omega B^r \left(Nx^r + q + K^r \left(x^{r+1} - x^r\right)\right)\right).$$

We also know that $\lambda > 0$. Using these and Theorem 2, we conclude that the first term in the right-hand side of the long equality above is ≤ 0. We therefore have that

$$\begin{aligned} f\left(x^{r+1}\right) - f\left(x^r\right) &\leq \left(x^{r+1} - x^r\right) \left(\tfrac{1}{2}N - \left(\lambda\omega B^r\right)^{-1} - K^r\right) \left(x^{r+1} - x^r\right) \\ &\leq -\beta \left\| x^{r+1} - x^r \right\|^2. \end{aligned} \tag{24}$$

The last inequality follows from conditions [A3] and proves the first part of our claim.

Since $\beta > 0$, the equation (24) implies that $f\left(x^r\right) - f\left(x^{r+1}\right) \geq 0$. Hence $\{f\left(x^r\right) : r = 1, 2, \ldots\}$ is a monotone nonincreasing sequence of real numbers. Let $\bar{x}$ be an accumulation point of the sequence $\{x^r : r = 1, 2, \ldots\}$. Hence, there exists a sequence of positive integers such that the sequence of x^r with r belonging to this subsequence of integers converges to $\bar{x}$. Since the sequences of $B^r$and $K^r$are bounded sequences of matrices, we can again find a subsequence of the above sequence of positive integers satisfying the property that both the subsequences of $B^r$and K^r with r belonging to this subsequence

converge to the limits. Let $\{r_t : t = 1, 2, \ldots\}$ be this final subsequence of positive integers. Therefore, we can assume that

$$\lim_{t\to\infty} B^{r_t} = \bar{B}, \ \lim_{t\to\infty} K^{r_t} = \bar{K}, \ \lim_{t\to\infty} x^{r_t} = \bar{x}.$$

In addition, from property (A2) it follows that $\bar{B}$ is a diagonal matrix with positive diagonal entries. Since $f(x)$ is a continuous function, we have that $f(\bar{x}) = \lim_{t\to\infty} f(x^{r_t})$. Since $\{f(x^r) : r = 0, 1, \ldots\}$ is a nonincreasing sequence of real numbers with a convergent subsequence, $f(x^{r_t})$, it follows that the entire sequence is itself convergent. This and (24) together imply that $0 = \lim_{t\to\infty} \left(f(x^{r_t}) - f\left(x^{r_t+1}\right)\right) \geq \lim_{t\to\infty} \beta \left\|x^{r_t} - x^{r_t+1}\right\|^2 \geq 0$. From this and the fact that the sequence $\{x^{r_t}\}$ is convergent to $\bar{x}$, it follows that $\left\{x^{r_t+1}\right\}$ is also convergent to the same limit. These facts imply that

$$\begin{aligned} 0 &= \lim_{t\to\infty} \left\|x^{1+r_t} - x^{r_t}\right\| \\ &= \lambda \left\|\Pi_{\Upsilon}\left(x^{r_t} - \omega B^{r_t}\left(N x^{r_t} + q + K^{r_t}\right]\left(x^{r_t+1} - x^{r_t}\right)\right)\right) - x^{r_t}\right\| \\ &= \lambda \left\|\Pi_{\Upsilon}\left(\bar{x} - \omega \bar{B}\left(N\bar{x} + q\right)\right) - \bar{x}\right\|. \end{aligned}$$

So we have that $\Pi_{\Upsilon}\left(\bar{x} - \omega\bar{B}\left(N\bar{x} + q\right)\right) - \bar{x} = 0$. By Theorem 1, we have that $(M\bar{x} + q, \bar{x})$ is a solution for CCP. □

Note that the preceding result does not mean that the sequence will converge, since it is still possible that the sequence will diverge to infinity and have no accumulation point. The proper alternative is related by the following result.

Theorem 4 *Under the assumptions of the section, either (a) the sequence x^r is bounded, or (b) there exists a $0 \neq y \in \Upsilon$, that satisfies $Ny = 0$. In case (a), any two accumulation points z^1 and z^2 satisfy $Nz^1 = Nz^2$.*

Proof Assume that case (a) does not hold. Then the sequence x^r will have a subsequence x^{r_i} that satisfies $x^{r_i} \to \infty$, as $i \to \infty$. Consider the sequence

$$y^i = \frac{x^{r_i}}{\|x^{r_i}\|},$$

which, being bounded, must have an accumulation point $\bar{y}$. We assume, without loss of generality, that the entire sequence y^i converges to $\bar{y} \in \Upsilon$. From the previous theorem, we have that the sequence $f(x^{r_i})$ is decreasing, and we obtain that

$$\frac{f(x^{r_i})}{\|x^{r_i}\|^2} = \frac{1}{2}\left(\frac{x^{r_i}}{\|x^{r_i}\|}\right)^T N \left(\frac{x^{r_i}}{\|x^{r_i}\|}\right) + \left(\frac{q^T}{\|x^{r_i}\|}\right)\left(\frac{x^{r_i}}{\|x^{r_i}\|}\right).$$

Taking the limit as $i \to \infty$, we obtain that $\bar{y}^T N \bar{y} \leq 0$. Using assumption [A1], we obtain that $\bar{y}^T N \bar{y} = 0$. From assumption [A1] it follows that $\bar{y}$ is a minimum for the function $y^T N y$, and from the optimality conditions it follows that $N\bar{y} = 0$. Since $\bar{y} \in \Upsilon$, case (b) must hold. We have thus proved that the outcome of the iterative method can be only (a) or (b).

Assume now that we are in case (a), and we have two accumulation points $z^1 \in \Upsilon$ and $z^2 \in \Upsilon$. From Theorem 3 we have that both z^1 and z^2 are a solution of (CCP), and therefore they satisfy the following relationships.

$$z^1 \in \Upsilon, z^2 \in \Upsilon, -(Nz^1 + q) \in \Upsilon^\circ, -(Nz^2 + q) \in \Upsilon^\circ.$$

Since both the cone Υ and the cone Υ° are convex sets, it follows that $z^2 + \lambda\left(z^1 - z^2\right) \in \Upsilon$ and that $Nz^2+q+\lambda\left(Nz^1 - Nz^2\right) \in -\Upsilon^\circ$, for any parameter $\lambda \in [0,1]$. In turn, from the definition of the polar cone, it follows that

$$g(\lambda) = \left\langle z^2 + \lambda\left(z^1 - z^2\right), Nz^2 + q + \lambda\left(Nz^1 - Nz^2\right)\right\rangle \geq 0, \quad \forall \lambda \in [0,1].$$

Using the fact that z^2 satisfies $\left\langle z^2, Nz^2 + q\right\rangle = 0$ being a solution of (CCP), we obtain that

$$\lambda\left\langle\left(z^1 - z^2\right), Nz^2 + q\right\rangle + \lambda\left\langle z^2, \left(Nz^1 - Nz^2\right)\right\rangle + \lambda^2\left\langle\left(z^1 - z^2\right), \left(Nz^1 - Nz^2\right)\right\rangle \geq 0, \quad \forall \lambda \in [0,1]$$

from which it follows that

$$\left\langle\left(z^1 - z^2\right), Nz^2 + q\right\rangle + \left\langle z^2, \left(Nz^1 - Nz^2\right)\right\rangle \geq 0.$$

By using the same argument but switching z^1 and z^2, we obtain that

$$-\left\langle\left(z^1 - z^2\right), Nz^1 + q\right\rangle - \left\langle z^1, \left(Nz^1 - Nz^2\right)\right\rangle \geq 0.$$

Adding the last two equations, we obtain that

$$\left\langle\left(z^1 - z^2\right), Nz^2 - Nz^1\right\rangle + \left\langle z^2 - z^1, \left(Nz^1 - Nz^2\right)\right\rangle \geq 0,$$

with which, using the symmetry of the matrix N that follows from assumption A1, we obtain that

$$2\left\langle\left(z^1 - z^2\right), \left(Nz^1 - Nz^2\right)\right\rangle \leq 0.$$

Using again assumption A1, we have that $Nz^1 = Nz^2$, which completes the proof. □

Corollary 1 *Assume that the friction cone of the configuration is pointed (that is, there does not exist a choice of reaction forces whose net effect is zero). If the relevant parameters satisfy assumptions [A2] and [A3], then the algorithm (21) for CCP applied to (19) produces a bounded sequence, and any accumulation point results in the same velocity solution.*

Proof Assume that the sequence x^r is produced by the algorithm (21) whose parameters satisfy Assumptions [A2] and [A3]. Then, from Theorem 4 there exists $0 \neq y \in \Upsilon$ such that $Ny = 0$. From (18) this implies that $Dy = 0$. In turn, from the definition of D in (18) and Subsection 2.3, this implies that there exist nonzero constraint feasible impulses that produce a zero net effect on the system. This contradicts the assumption that the friction cone is pointed [38].

Therefore boundedness of the iteration sequence and the existence of an accumulation point are assured. Uniqueness of the velocity follows from the second part of Theorem 4, since $Nz_1 = Nz_2$ and the definition of D in (18) implies that $Dz_1 = Dz_2$, which, in turn, from (16) implies that the velocity solution is unique. The proof is complete. □

Note that N is symmetric positive semi-definite and therefore assumption [A1] is satisfied. Assumption [A2] is easily satisfied, whereas assumption [A3] can be satisfied by a trial-and-error approach, whereas if the iterates x^{r+1} and x^r do not satisfy [A3], than the parameter ω is decreased by factor of two, and the iterate $x^{(r+1)}$ is recomputed. It is immediate that such a strategy can decrease the parameter ω only a finite number of times.

4 Implementation

The CCP method proposed here can be applied to the simulation of multibody systems with a large number of parts and contacts because, where an upper limit on the number of iteration is enforced, the iteration (21) can run in $O(n)$ space and $O(n)$ time.

Previous sections showed that generic multibody problems with frictional contacts, expressed with the system (5)–(8), embed the cone complementarity problem (19). Hence, the iterative method (21) can be used to solve such convex CCP, because (19) is equivalent to the more general problem (20) where one considers the specific case of three-dimensional cones Υ_i. That is, for the ith friction cone Υ_i we have that $n_i = 3$ and that there is an associated vector with a normal and two tangential reactions: $\gamma^i = \{\gamma^i_n, \gamma^i_u, \gamma^i_v\}$. The complete vector of unknown scalar reactions is $\gamma \in \mathbb{R}^{3n_\mathcal{A}}$. From Section 3, we have that $n_k = n_\mathcal{A}$ and $n_c = 3n_\mathcal{A}$.

Given (17), the final time-stepping scheme can be seen as a sequence of three main operations: a CCP problem that finds unknown reactions γ (25a), a linear application (25b) that gives the new speeds $v^{(l+1)}$, and a position update (25c):

$$(N\gamma + d) \in -\Upsilon^o \quad \perp \quad \gamma \in \Upsilon \tag{25a}$$

$$v^{(l+1)} = M^{-1}\left(\tilde{k} + D\gamma\right) \tag{25b}$$

$$q^{(l+1)} = q^{(l)} + h\Gamma(q^{(l)})v^{(l+1)}. \tag{25c}$$

The biggest computational overhead is caused by the first problem, that is, the CCP (25a). In fact, (25c) is immediate, and (25b) can be computed quickly because in most cases the matrix M is diagonal and its inverse M^{-1} can be precomputed easily.

We recall that $N = D^T M^{-1} D$. The full D matrix can be partitioned in $n_\mathcal{A}$ vertical blocks $D_i \in \mathbb{R}^{m\times 3}$, each pertaining to the corresponding ith cone. We also recall that, from (18), we have that

$$D = \left[D^1|D^2|...|D^{n_\mathcal{A}}\right].$$

Using (17), we can rewrite the term r from (17) in a more compact form:

$$d = D^T M^{-1}\tilde{k} + b. \tag{26}$$

The convergence theory about the iterative scheme (21) leaves some degrees of freedom in choosing η_i values that build the diagonal blocks of the

iteration matrix B. A trivial choice could be to use the same $\eta_i = \xi$ value for all diagonal blocks, that is, $B = \xi I$, and then use the overrelaxation parameter ω to control the convergence. However, setting the same value for all η_i may slow convergence in systems with large mass ratios, even with an optimal ω. A more practical approach, which copes better with systems affected by uneven masses, is to use $\eta_i = 1/\bar{g}_i$, where $\bar{g}_i$ is the average of the diagonal values of the ith block of the N matrix. We note that $\bar{g}_i$ can be computed easily from the trace of the 3×3 matrix $D^{i,T} M^{-1} D^i$, as

$$\bar{g}_i = \frac{\mathrm{Trace}(D^{i,T} M^{-1} D^i)}{3}. \tag{27}$$

Also the K matrix in (21) can be chosen freely, within the convergence limits posed by assumptions [A1]–[A3]. Among the most noticeable options, we note the case where $K = 0$, which results in a scheme like a projected Jacobi, or the case where K is built by using the lower blocks of N, so that $K_{i,j} = D^{i,T} M^{-1} D^j$, where $1 \leq j < i \leq n_{\mathcal{A}}$. In practical terms this means that, as soon as computed, a triplet γ^i with three reaction values will be used also for computing the following γ^{i+1} triplet, and so on for all i, without needing to finish a single iteration, which results in a Gauss-Seidel-type iteration. Numerical tests show that this last option, similar to a projected SOR scheme with immediate update of unknown vector, converges faster than the case of $K = 0$. Hereafter, we will assume that such a kind of K matrix is used. Another choice, that we do not explore here, is the one of having the matrix K be block-diagonal with block lower triangular blocks, which would be equivalent with block Jacobi, where for each block we do a Gauss-Seidel-type iteration. The latter is suitable for a parallel iteration with low communication overhead.

We recall that the matrix N is a product of large matrices; $N = D^T M^{-1} D$, and it is full even if D and M are sparse. For systems with a large number of contacts, the size of N would be prohibitive and clearly would not satisfy the goal of $O(n)$ space complexity. To this end, direct multiplication of vectors and matrices in (21) must be avoided; otherwise the effort and the space requirement would be superlinear in the number of constraint.

For the reasons above, a scheme that does not need N computed explicitly has been developed, exploiting the sparsity of M and D. Also the K matrix does not need to be explicitly built, if we adopt the above mentioned choice of K as the upper block-structure of N. These considerations lead to the following implementation of the rth step of the iteration (21), expressed as an inner loop with index $i = 1 \ldots n_{\mathcal{A}}$ on all $n_{\mathcal{A}}$ friction cones Υ^i:

$$\begin{aligned} \delta^{ir} = \gamma^{ir} - \omega\eta_i \Bigg(D^{i,T} M^{-1} \Bigg(\sum_{z=1}^{i-1} D^z \gamma^{z,r+1} + \\ + \sum_{z=i}^{n_{\mathcal{A}}} D^z \gamma^{zr} + \tilde{k}^i \Bigg) + b^i \Bigg) \end{aligned} \tag{28}$$

$$\gamma^{i,r+1} = \lambda \Pi_{\Upsilon^i} \left(\delta^{ir} \right) + (1 - \lambda) \gamma^{ir}. \tag{29}$$

In the case of friction in three-dimensional space, the implementation of the projection operator $\Pi_{\Upsilon^i}(\delta^i) : \mathbb{R}^3 \to \mathbb{R}^3$ is straightforward.

For improved performance, some operations can be computed at the beginning of the iteration because their outcome would remain unchanged. In detail, we introduce the $m \times 3$ matrix $s^i = M^{-1}D^i$ and the 3×3 matrix $g^i = D^{i,T}M^{-1}D_i$.

Considering the optimizations above, we can express the final CCP algorithm with the following pseudocode:

Algorithm 1

1. For $i = 1, 2, \ldots, n_{\mathcal{A}}$ compute the $m \times 3$ matrices $s^i = M^{-1}D^i$ and 3×3 matrices $g^i = D^{i,T}s^i$.
2. For $i = 1, 2, \ldots, n_{\mathcal{A}}$, compute $\eta_i = 3/\mathrm{Trace}(g^i)$.
3. If warm starting with some initial guess γ^*, initialize reactions as $\gamma^0 = \gamma^*$, otherwise $\gamma^0 = 0$.
4. Initialize speeds: $v = \sum_{i=1}^{n_{\mathcal{A}}} s^i\gamma^0 + M^{-1}\tilde{k}$.
5. For $i = 1, 2, \ldots n_{\mathcal{A}}$, perform the updates
   $\delta^{i,r} = \left(\gamma^{i,r} - \omega\eta_i\left(D^{i,T}v^r + b^i\right)\right)$;
   $\gamma^{i,r+1} = \lambda\Pi_{\Upsilon}\left(\delta^{i,r}\right) + (1-\lambda)\gamma^{i,r}$;
   $\Delta\gamma^{i,r+1} = \gamma^{i,r+1} - \gamma^{i,r}$;
   $v := v + {s^i}^T\Delta\gamma^{i,r+1}$.
6. Repeat the step 5 by looping on the list of contacts in backward direction, if symmetric updates are desired.
7. $r := r + 1$. Repeat from 5 until convergence, or until $r > r_{max}$.

The iterations, usually stopped when an approximation threshold has been reached, can be also prematurely aborted when r exceeds a limit r_{max} on the maximum number of iterations if the simulation must meet hard-real-time requirements.

With minimal changes to the $\Pi_{\Upsilon}(\cdot)$ operator, the proposed method can easily adapted to the case of friction in 2D or the case of generic unilateral constraints. Also, without major modifications to the main scheme, classical bilateral constraints can be added. In that case, the cones Υ^i can be taken to be $\mathbb{R}^{l_i}$, where l_i is the dimension of the force vector associated to a joint and the projection on such a cone is the identity. Therefore the scheme is modified by simply replacing (29) with

$$\gamma^{i,r+1} = \lambda\delta^{ir} + (1-\lambda)\gamma^{ir}. \tag{30}$$

In our simulations, we chose $\omega = 1$ and $\lambda = 1$, except for the test of convergence of the residual with respect to ω. We cannot guarantee a priori that this will satisfy condition [A3], but it did for all our simulations. In addition, the matrix sequences K^r and B^r were constant. We can therefore claim that Theorem 3 does apply and, since the sequence did not diverge (and was in fact convergent), any accumulation point is a solution of the cone complementarity problem (25a). In addition, Theorem 4 is applicable

to show that any accumulation point has the same velocity solution. It is difficult to verify numerically the condition of Corollary 1 (the pointed friction cone assumption). Nonetheless, boundedness of the iterates was observed in all cases. In addition, our proofs of the theoretical results allow for similar conclusions if ω varies from iteration to iteration. Therefore, we could ensure that at some iteration the appropriate ω is chosen after decreasing its value a few times until assumption [A3] holds. It can be shown that if the value of ω is halved each time [A3] does not hold and the respective iteration is rejected, then [A3] will eventually be satisfied after a finite number of steps. In our experiments, however, the values we have chosen for ω and λ have worked for all iterations without need of further adjustment.

Finally, the overall scheme (25) for advancing a single time step can be expressed with the following pseudocode:

Algorithm 2

1. Set $t = 0$, step counter $l = 0$, provide initial values for $v^{(l)}$ and $q^{(l)}$.
2. Perform collision detection between shapes of bodies, obtaining $n_{\mathcal{A}}$ possible contact points within an ϵ distance. For each contact point, compute D_n, D_u, D_v and residual $\Phi(q)$.
3. For each body, compute forces $f_t(q, v, t)$, then compute $\tilde{k}$ with (15) and b with (17).
4. Use Algorithm 1 to obtain unknown impulses γ and speeds $v^{(l+1)}$ for the CCP problem.
5. Update positions using $q^{(l+1)} = q^{(l)} + h\Gamma(q^{(l)})v^{(l+1)}$.
6. Increment $t := t + h$, $l := l + 1$, and repeat from step 2 until $t > t_{\text{end}}$

We remark that choosing a proper value for the collision envelope ϵ is not trivial. If a very small or zero value is used, contacts will enter the CCP solver only when it is too late and some amount of interpenetration will be unavoidable: this will affect negatively the stability of the method. On the other hand, if too large values are used, the collision detection algorithm will return too many potential contacts which will waste computational resources for the CCP solution and which could occasionally create troubles with convex shapes: this will decrease the efficiency and the robustness of the method. We experienced that a simple yet efficient heuristic is to choose ϵ as the maximum distance which can be spanned by whatever point of the moving body, given its speed.

5 Examples

We present the results of our algorithm on two granular materials applications. For the larger simulation, the number of impulse variables exceeded 400,000.

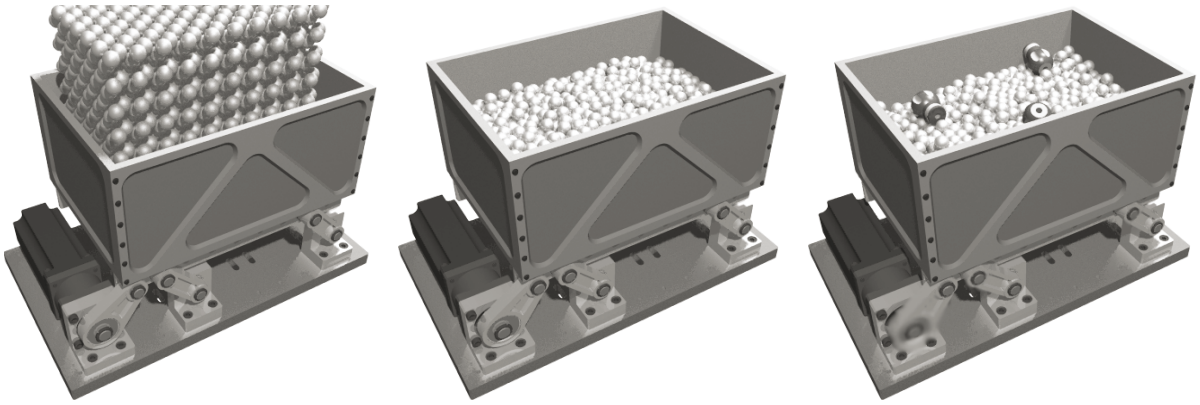

Fig. 2 Shaker benchmark, for a 1500-sphere case, showing vibration-induced granular segregation of large objects. After ten seconds of simulation three black intruder particles, originally placed at the bottom of the shaker, will rise at the top surface.

5.1 Size segregation in a shaker

The first example is meant as a benchmark to evaluate the performance of the solver when dealing with many contacts with friction. A rectangular box is filled with spheres; then the box is shaken by means of an articulated suspension and a crank mechanism (Fig. 2). When large objects are mixed with the spheres, a phenomenon called vibration-induced size segregation moves larger objects on top: this effect can be observed also in our simulations.

Different parameters have been tested, for example repeating the simulation with a varying number of spheres up to 1500. In all cases the mass of the spheres is $m = 0.01$kg, their diameter is $d = 26$mm, the friction coefficient is $\mu = 0.3$, and the time step is $h = 2\pi/50\Omega$s, with Ω rad/s being the frequency of the crank. The amplitude of the vibration has been tested up to $A = 10$mm.

Plotting of $||\Delta\gamma||$ (Fig. 3) during the iteration of the algorithm shows the convergence of the method for varying values of the overrelaxation factor ω.

Figure 4 shows how the CPU spends time in various parts of the simulation algorithm. For this benchmark, a shaker with 1,000 rigid bodies was simulated with an upper limit of 40 iterations for the CCP solver. One can easily see that the solution of the CCP is the bottleneck in the entire simulation process, while collision detection and other tasks (time integration, Jacobian update, etc.) are less CPU-intensive. In this example, the 40-iteration limit was enough to keep the feasibility errors at negligible levels (max. interpenetration $||\epsilon_{Pn}|| < 0.002\,d$). However, if lower precision is acceptable as in case of virtual reality or real-time applications, fewer iterations can be used, thus reducing the CCP timings to levels which are comparable to the collision detection timings.

To show how the number of iterations affect the precision of the solution to the complementarity problem, in Fig. 5 we report the maximum error in terms of speed violation $||\epsilon_{Vn}||$ in contact constraints, during 300 time steps of simulation. Speed is measured in d/s, where d is the diameter of the spheres.

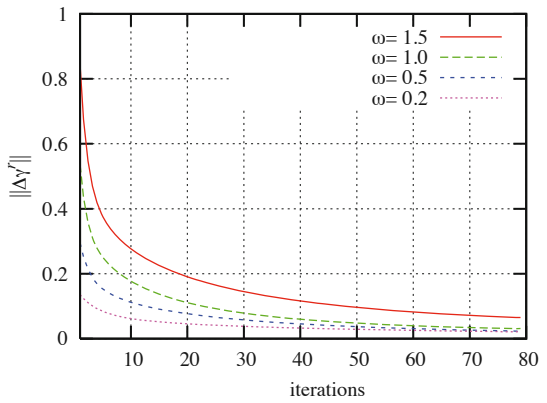

Fig. 3 Convergence of $\Delta\gamma^r$ for varying ω, for a sample time step in the 300-sphere benchmark.

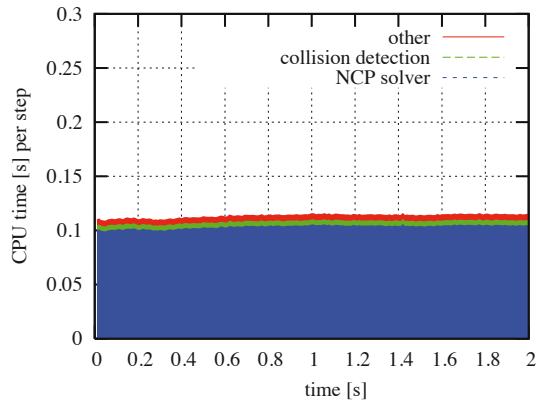

Fig. 4 CPU time for each step in a 1000-body simulation, split into CCP fraction, collision detection fraction, and other.

Similarly, we report in Fig. 6 the maximum position error $||\epsilon_{Pn}||$ in contact constraints, that is, the maximum interpenetration. The error is measured in d units. One can see that, despite the large number of objects in contact, acceptable precision can be obtained also with a moderate number of iterations.

By performing a set of six shaker simulations with an increasing number of objects, hence for increasing numbers of contacts, one can obtain a graph as in Fig. 8, which shows how the CPU effort grows linearly with the number of frictional contacts. Here, for each simulation, CCP timings have been recorded after ten seconds of transient, when spheres are at steady state and form a dense packing, because this is a nontrivial configuration that requires significant CPU efforts. The linear-time complexity is a consequence of the loop in the fifth step of the algorithm, which is $O(p)$ with p reaction forces γ_i if a maximum number of iterations is enforced (40 iterations in this example). Note that the fourth step of the algorithm, performing a computation that is linear in terms of number of rigid bodies, has a moderate or negligible impact on overall performance. Despite the fact that the theoretical complexity of the algorithm for fixed number of iterations is linear, some deviation from linearity can be experienced in complex applications when large amounts of contacts are simulated, because CPU cache misses can become more frequent as the memory access starts to become more and more intense.

Figure 7 show that only a portion of the potential contacts will be active (i.e., with nonzero reaction force) after the CCP solution. Since the computational effort is proportional to the number of potential contacts entering the CCP solver, regardless of the active/inactive outcome, a proper collision detection algorithm should take care to report the smallest number of potential contacts, that is, only the surface pairs that may give interpenetration in a single time step given the actual state of bodies. This precaution would keep the active contacts as a relatively fixed percentage of the number of potential constraints.

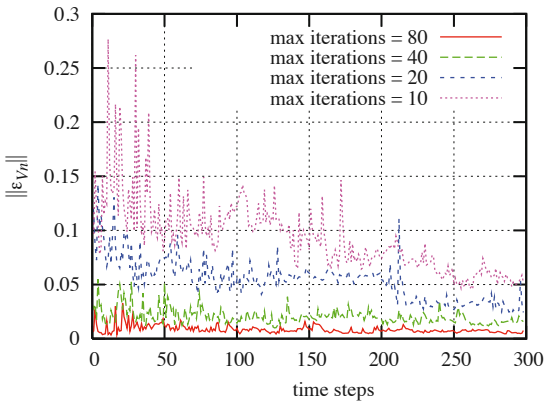

Fig. 5 Maximum speed violation in constraints, for the 300-sphere benchmark.

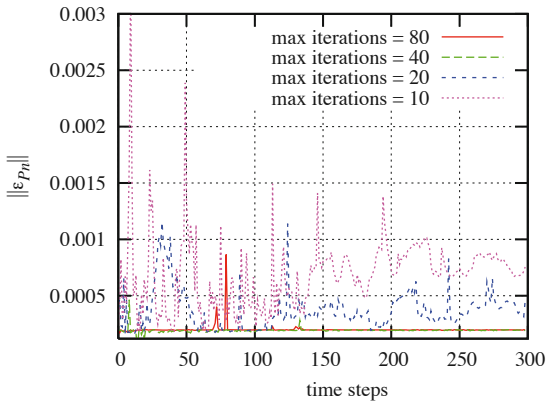

Fig. 6 Maximum penetration error in constraints, for the 300-sphere benchmark.

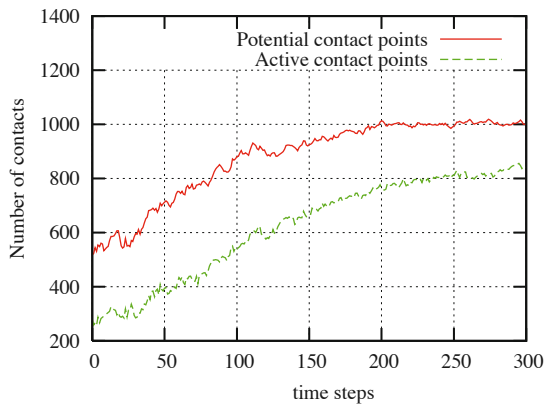

Fig. 7 Number of contact constraints, increasing while pouring spheres in the shaker.

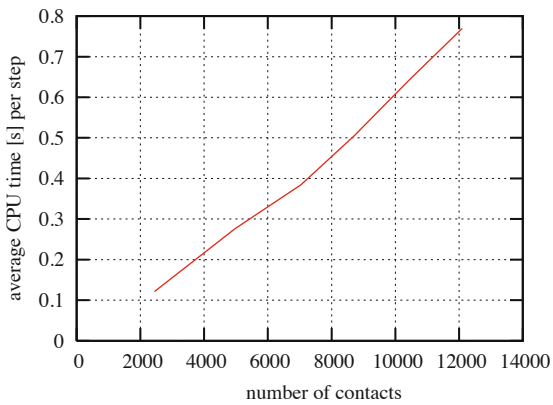

Fig. 8 Average CPU time used to compute a step of simulation, as a function of the number of contacts.

5.2 Granular flow from a silo

The numerical method proposed in this article can be used to simulate dense granular flows in silos. This problem arises in many engineering applications, most noticeably in the development of the promising fourth-generation uranium-based, graphite-moderated, helium-cooled very high temperature

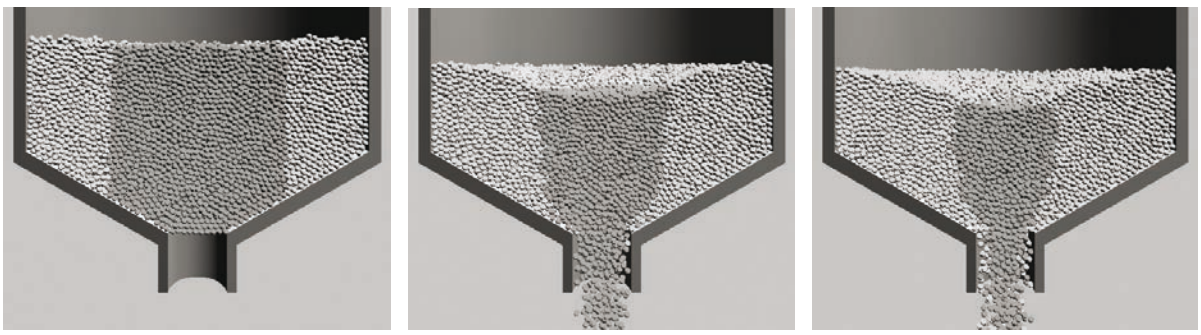

Fig. 9 Frames from the simulation of 36,000 rigid bodies with frictional contacts flowing from a three-dimensional funnel. On a T2600 2 GHz processor, each solution of the CCP problem (nearly half a million of variables, including constraint multipliers and speeds) with 140 iterations, took 19 s of CPU time on average.

nuclear reactor, where thousands of graphite fuel pebbles drain very slowly in a continuous refueling process [13].

Pebble flow in such pebble-bed reactors (PBRs) is not easily accessible to experiments, and no reliable continuum model is yet available for analytical approaches. These facts motivate the development of fast numerical methods. Simulations of PBR reactors have been recently performed with DEM discrete-element methods [33]; however, the DEM approach is based on a stiff spring-dashpot contact model which requires a very small time step in order to guarantee the stability of the integration. Conversely, the method proposed here can enforce rigid contacts without the need of artificial stiffness; hence larger timesteps can be used. For example, the flow simulation of Fig.9, representing 11 seconds of drainage from a silo 3.5m wide with 36,000 uranium-graphite spheres with $d = 0.06$m and friction coefficient $\mu = 0.6$, exploited a timestep $h = 0.01$s that is three orders of magnitude larger than the timestep required by the DEM method in [33]. The simulation took about four hours to complete for 5 seconds of simulated time, with a penetration error comparable to the one in the size segregation case. But timing is perhaps less relevant since it depends on items such as cache management that can vastly change with different optimization than the fact that the simulation completed with low penetration error for a fixed (and relatively small) number of iterations, 140, for a very high density configuration. The maximum number of contacts for which the problem was benchmarked was almost half a million, which in turn resulted in more than two millions of variables for the CCP. This is a promising approach to the simulation of full-scale reactors and other large granular flow problems, though further tests are needed to determine whether the maximum penetration error does not increase with an increasing number of uranium-graphite spheres.

6 Conclusions

Aiming at a linear-time solution of dynamical systems with thousands of constraints and contacts, we have presented a novel method for solving the cone constrained subproblems that appear in a time-stepping approach recently proposed in [1]. The method has the flavor of a Gauss-Seidel with overrelaxation and is proven to converge under fairly standard assumptions about the configuration of the system.

We implemented this method into the HyperOctant library of our multibody project, Chrono::Engine [39]. Our method is able to handle large simulations with tens of thousands of colliding rigid bodies and hundreds of thousands of constraint impulse variables and scales well in this range. In previous work [4, 40] we have shown that simplex-like methods do not scale well for systems in configurations of the type solved here. In future work, and as appropriate software packages become available, we will carry out comparisons with interior-point methods for optimization problems with conic constraints.

Because of the low computational overhead of our method, we foresee that it could be endorsed even in the emerging application fields of physical engines

for videogames and virtual interactive environments, which can exploit the benefits of the method for real-time performance.

Acknowledgments

We thank Paul Tseng for technical discussions concerning block coordinate descent methods. We are profoundly grateful to the two referees and to Dan Negrut for their constructive comments. Mihai Anitescu was supported by Contract No. DE-AC02-06CH11357 of the U.S. Department of Energy.

References

1. M. Anitescu. Optimization-based simulation of nonsmooth rigid multibody dynamics. *Math. Program.*, 105(1):113–143, 2006.
2. M. Anitescu, J. F. Cremer, and F. A. Potra. Formulating 3d contact dynamics problems. *Mechanics of Structures and Machines*, 24(4):405–437, 1996.
3. M. Anitescu and G. D. Hart. A constraint-stabilized time-stepping approach for rigid multibody dynamics with joints, contact and friction. *International Journal for Numerical Methods in Engineering*, 60(14):2335–2371, 2004.
4. M. Anitescu and G. D. Hart. A fixed-point iteration approach for multibody dynamics with contact and friction. *Mathematical Programming, Series B*, 101(1):3–32, 2004.
5. M. Anitescu and F. A. Potra. Formulating dynamic multi-rigid-body contact problems with friction as solvable linear complementarity problems. *Nonlinear Dynamics*, 14:231–247, 1997.
6. M. Anitescu and F. A. Potra. Time-stepping schemes for stiff multi-rigid-body dynamics with contact and friction. *International Journal for Numerical Methods in Engineering*, 55(7):753–784, 2002.
7. M. Anitescu, F. A. Potra, and D. Stewart. Time-stepping for three-dimensional rigid-body dynamics. *Computer Methods in Applied Mechanics and Engineering*, 177:183–197, 1999.
8. D. Baraff. Issues in computing contact forces for non-penetrating rigid bodies. *Algorithmica*, 10:292–352, 1993.
9. D. Baraff. Fast contact force computation for nonpenetrating rigid bodies. In *Computer Graphics (Proceedings of SIGGRAPH)*, pages 23–34, 1994.
10. R. Cottle and G. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1:103–125, 1968.
11. R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, Boston, 1992.
12. B. R. Donald and D. K. Pai. On the motion of compliantly connected rigid bodies in contact: a system for analyzing designs for assembly. In *Proceedings of the Conf. on Robotics and Automation*, pages 1756–1762. IEEE, 1990.
13. H. D. Gougar. *Advanced core design and fuel management for pebble-bed reactors.* Ph.D thesis, Department of Nuclear Engineering, Penn State University, 2004.
14. E. J. Haug. *Computer Aided Kinematics and Dynamics of Mechanical Systems.* Allyn and Bacon, Boston, 1989.
15. E. J. Haug, S. Wu, and S. Yang. Dynamic mechanical systems with coulomb friction, stiction, impact and constraint addition-deletion. *Mechanisms and Machine Theory*, 21(5):407–416, 1986.
16. J.-B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms.* Springer Verlag, Berlin, 1993.
17. F. Jourdan, P. Alart, and M. Jean. A Gauss Seidel like algorithm to solve frictional contract problems. *Computer methods in applied mechanics and engineering*, 155:31 –47, 1998.

18. Y. J. Kim, M. C. Lin, and D. Manocha. Deep: Dual-space expansion for estimating penetration depth between convex polytopes. In *Proceedings of the 2002 International Conference on Robotics and Automation*, volume 1, pages 921–926. Institute for Electrical and Electronics Engineering, 2002.
19. P. Lotstedt. Mechanical systems of rigid bodies subject to unilateral constraints. *SIAM Journal of Applied Mathematics*, 42(2):281–296, 1982.
20. M. D. P. Marques. *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, volume 9 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser Verlag, Basel, Boston, Berlin, 1993.
21. J. J. Moreau. Standard inelastic shocks and the dynamics of unilateral constraints. In G. D. Piero and F. Macieri, editors, *Unilateral Problems in Structural Analysis*, pages 173–221, New York, 1983. CISM Courses and Lectures no. 288.
22. J. J. Moreau and M. Jean. Numerical treatment of contact and friction: The contact dynamics method. In *Proceedings of the Third Biennial Joint Conference on Engineering Systems and Analysis*, pages 201–208, Montpellier, France, July 1996.
23. R. M. Murray, Z. Li, and S. S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Boca Raton, FL, 1993.
24. K. G. Murty. *Linear Complementarity, Linear and Nonlinear Programming*. Helderman Verlag, Berlin, 1988.
25. J.-S. Pang, V. Kumar, and P. Song. Convergence of time-stepping method for initial and boundary-value frictional compliant contact problems. *SIAM J. Numer. Anal.*, 43(5):2200–2226, 2005.
26. J.-S. Pang, V. Kumar, and J. Trinkle. On a continuous-time quasistatic frictional contact model with local compliance. *International Journal for Numerical Methods in Engineering*, 2007. submitted.
27. J.-S. Pang and D. Stewart. A unified approach to frictional contact problems. *Internat. J. Eng. Science*, 37(13):1747–1768, 1999.
28. J.-S. Pang and D. Stewart. Solution dependence on initial conditions in differential variational inequalities. *Set-Valued Analysis*, page submitted, 2004.
29. J.-S. Pang and D. Stewart. Differential variational inequalities. *Mathematical Programming*, 113(2):345–424, 2008.
30. J.-S. Pang and J. C. Trinkle. Complementarity formulations and existence of solutions of dynamic multi-rigid-body contact problems with coulomb friction. *Math. Program.*, 73(2):199–226, 1996.
31. F. Pfeiffer and C. Glocker. *Multibody Dynamics with Unilateral Contacts*. John Wiley, 1996.
32. F. A. Potra, M. Anitescu, B. Gavrea, and J. Trinkle. A linearly implicit trapezoidal method for integrating stiff multibody dynamics with contact and friction. *International Journal for Numerical Methods in Engineering*, 66(7):1079–1124, 2006.
33. C. Rycroft, G. Grest, J. Landry, and M. Bazant. Analysis of granular flow in a pebble-bed nuclear reactor. *Physical Review E*, 74, 021306, 2006.
34. P. Song, P. Kraus, V. Kumar, and P. Dupont. Analysis of rigid-body dynamic models for simulation of systems with frictional contacts. *Journal of Applied Mechanics*, 68(1):118–128, 2001.
35. P. Song, J.-S. Pang, and V. Kumar. A semi-implicit time-stepping model for frictional compliant contact problems. *International Journal of Numerical Methods in Engineering*, 60(13):267–279, 2004.
36. D. E. Stewart. Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painleve's problems. *Archive Rational Mechanics and Analysis*, 145(3):215–260, 1998.
37. D. E. Stewart. Rigid-body dynamics with friction and impact. *SIAM Review*, 42(1):3–39, 2000.
38. D. E. Stewart and J. C. Trinkle. An implicit time-stepping scheme for rigid-body dynamics with inelastic collisions and Coulomb friction. *International Journal for Numerical Methods in Engineering*, 39:2673–2691, 1996.

39. A. Tasora. Chrono::engine project, web page. *www.deltaknowledge.com/chronoengine*, 2006.
40. A. Tasora, E. Manconi, and M. Silvestri. Un nuovo metodo del simplesso per il problema di complementarit lineare mista in sistemi multibody con vincoli unilateri. In *Proceedings of AIMETA 05*, Firenze, Italy, 2005.
41. J. Trinkle, J.-S. Pang, S. Sudarsky, and G. Lo. On dynamic multi-rigid-body contact problems with coulomb friction. *Zeithschrift fur Angewandte Mathematik und Mechanik*, 77:267–279, 1997.
42. P. Tseng and S. Yun. A coordinate gradient descent method for nonsmooth separable minimization. *Mathematical Programming B*, 2007. submitted.

A Notations

A.1 Multibody system

- n: number of bodies.
- $m = 6n$: dimension of the position state vector.
- M: Mass matrix, positive definite, of size $m \times m$.
- q: vector of generalized positions of dimension m.
- v: vector of generalized velocites of dimension m.
- t: time of the system.
- h: time step used by the time-stepping scheme.
- $f_t(t, q, v), f_e(t, q, v), f_c(q, v)$: the total, external, and, respectively, Coriolis forces acting on the system, vectors of dimension m.
- p: number of contact constraints that can become active (no more than $\begin{pmatrix} n \\ 2 \end{pmatrix}$).
- $\mathcal{A}(q^l, \epsilon)$ set of ϵ-active contact constraints, a set with no more than p elements.
- $n_{\mathcal{A}}$: dimension of set $\mathcal{A}(q^l, \epsilon)$.
- D: aggregate matrix of normal and tangential directions at the contact in generalized coordinates, a matrix of dimension $m \times 3n_{\mathcal{A}}$.

A.2 Contact and friction model

- $\Phi(q)$: the gap function, which indicates whether a contact constraint is active.
- $\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2$: the normal and tangential vectors at a contact, three-dimensional vectors.
- $\mathbf{F}_N, \mathbf{F}_T$: the normal and tangential force at a contact, three-dimensional vectors.
- $\mathbf{v}_T$: the tangential velocity, a three-dimensional vector.
- $\widehat{\gamma}_n, \widehat{\gamma}_u, \widehat{\gamma}_v$: normal and tangential force multipliers.
- $\gamma_n, \gamma_u, \gamma_v$: normal and tangential impulse multipliers.
- D_n, D_u, D_v: the normal and tangential vectors at a contact in generalized coordinates, m-dimensional vectors.
- F_N, F_T: the normal and tangential force in generalized coordinates, m-dimensional vectors.
- μ: the friction coefficient.
- $\mathcal{FC}$: the Coulomb friction cone, a subset of a three-dimensional space.

A.3 Cone Complementarity Problems

- n_k: number of cones Υ_i whose direct sum give the total constraint cone of the cone complementarity problem.
- n_i: dimension of the vector space $\mathbb{R}^{n_i}$ in which the component Υ_i is embedded.

- $n_c = \sum_{i=1}^{n_k} n_i$: dimension of the unknown vector of the cone complementarity problem.
- Υ: the total constraint cone of the cone complementarity problem, a subset of $\mathbb{R}^{n_c}$.
- N: matrix of the cone complementarity problem of dimension $n_c \times n_c$.
- d: free term of the cone complementarity problem, a vector of dimension n_c.
- C°: the polar cone of a convex cone C.
- $\Pi_C(\cdot)$: the projection operator on a closed, convex cone C.

A.4 Iterative scheme

- K: a strict block upper or lower triangular matrix.
- B: a block diagonal matrix, with multiple of identity blocks.
- ω, λ: scalar parameters of the iterative scheme.

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.