Meccanica (2013) 48:1643–1659
DOI 10.1007/s11012-013-9694-y

A complementarity-based rolling friction model for rigid contacts

Alessandro Tasora · Mihai Anitescu

Received: 26 August 2012 / Accepted: 9 January 2013 / Published online: 23 January 2013

Abstract In this work (also, preprint ANL/MCS-P3020-0812, Argonne National Laboratory) we introduce a complementarity-based rolling friction model to characterize dissipative phenomena at the interface between moving parts. Since the formulation is based on differential inclusions, the model fits well in the context of nonsmooth dynamics, and it does not require short integration timesteps. The method encompasses a rolling resistance limit for static cases, similar to what happens for sliding friction; this is a simple yet efficient approach to problems involving transitions from rolling to resting, and vice-versa. We propose a convex relaxation of the formulation in order to achieve algorithmic robustness and stability; moreover, we show the side effects of the convexification. A natural application of the model is the dynamics of granular materials, because of the high computational efficiency and the need for only a small set of parameters. In particular, when used as a micromechanical model for rolling resistance between granular particles, the model can provide an alternative way to capture the effect of irregular shapes. Other applications can be related to real-time simulations of rolling parts in bearing and guideways, as shown in examples.

Keywords Variational inequalities · Contacts · Rolling friction · Multibody · Complementarity

A. Tasora (✉)
Dipartimento di Ingegneria Industriale, Università degli Studi di Parma, 43100 Parma, Italy
e-mail: tasora@ied.unipr.it

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

1 Introduction

Rolling resistance has attracted the interest of researchers since the early ages of applied mechanics [8, 26, 37]. During the past two centuries many models have been proposed, depending on the required level of detail and on the type of phenomena that cause the rolling resistance; usually the number of parameters increases with the complexity of the model. At one end of the spectrum, for instance, there is case of the rolling tire, which often requires sophisticated models with a large number of parameters [24]. In this work, on the other hand, we are interested in a model that has a small number of parameters but is easily applicable to problems with large number of parts, or with requirements of high computational efficiency in general. Superior performance stems from the adoption of a set-valued formulation that finds the solution in terms of a complementarity problem.

The advantages of a complementarity-based rolling friction model are multiple: it encompasses both the

moving and the static cases, it does not require regularization and stiff force fields, it is an intuitive extension of the classic Coulomb-Amontons sliding friction model to the rolling case, and it requires few parameters. A literature search reveals only a few contributions on this topic; among these we cite the relevant work of [20].

A complementarity-based model can be particularly useful in simulations that require algorithmic robustness and efficiency, such as in real-time simulators, robotics, and virtual and augmented reality. To this end we present an application of the method to the efficient simulation of a linear guideway with recirculating ball bearings.

Another motivation, also discussed in this paper, is the simulation of rolling resistance among a large number of parts. This happens, for example, when studying granular materials, such as in the interaction between machines and soils (tracked vehicles on sand, tires on deformable ground, etc.) In the context of granular material dynamics, rolling friction between microscopic spherical particles has the side effect of approaching, on a macroscopic scale, the same global behavior of the granular assembly if it were modeled with many faceted irregular shapes: there exists a relation between the degree of irregularity of the surfaces and the rolling friction coefficient [9].

In the literature, typical approaches to the simulation of contacts are based on the regularization of contact forces; such forces, which are discontinuous in nature, are approximated by smooth functions. Smoothness allows the problem to be dealt as system of ordinary differential equations (ODEs), the drawback being that the resulting ODEs, while tractable, are stiff, thereby requiring short timesteps and leading to high computational times [12]. The adoption of implicit integrators might alleviate, but not eliminate, the issue of stiffness. For instance, since the seminal work of [7], most implementations of the discrete element method (DEM) for the simulation of granular materials are based on regularization of contact forces and stiff ODEs. For very large systems, regardless of the fact that one can exploit powerful hardware, the computational time can be so high that some problems become untractable.

For this reason in [32] we proposed an approach based on differential variational inequalities (DVIs), as an alternative to the classical regularization-based approaches. The DVI approach, whose capabilities are not necessarily confined to granular problems, is a recent general way of dealing with nonsmooth mechanical problems; the approach encompasses ODEs as subcases as well as complementarity-based methods. In DVIs one can describe forces by means of set-valued functions (multifunctions) that capture the nonsmooth nature of models such as the Signorini contact law or the Coulomb friction; no stiff regularizations of discontinuities are needed because discontinuities are presented directly as complementarity constraints. Large timesteps are allowed, but at the cost of solving a variational inequality problem (a complementarity problem, in the simplest case) for each timestep [25].

A generic variational inequality (VI) is a problem of the type

$$\mathbf{u} \in \mathbb{K}: \quad \big\langle F(\mathbf{u}), \mathbf{y} - \mathbf{u} \big\rangle \geq 0 \quad \forall \mathbf{y} \in \mathbb{K} \tag{1}$$

given a closed and convex $\mathbb{K} \in \mathbb{R}^n$ set, and given a continuous $F(\mathbf{u}) : \mathbb{K} \to \mathbb{R}^n$. We call $\mathsf{SOL}(\mathbb{K}, F)$ the solution of problem (1). Variational inequalities are powerful mathematical tools that have recently been used also in game theory, continuum mechanics and other scientific fields; a good reference is [18].

Assuming that the state of the system is defined by $\mathbf{x}$, one can define a DVI as the problem of finding the function $\mathbf{x}$ on $[0, T]$,

$$\frac{d\mathbf{x}}{dt} = f(t, \mathbf{x}, \mathbf{u}) \tag{2}$$

$$\mathbf{u} \in \mathsf{SOL}\big(\mathbb{K}, F\big(t, \mathbf{x}(t), \cdot\big)\big), \tag{3}$$

along with boundary conditions $\Gamma(\mathbf{x}(0), \mathbf{x}(T)) = 0$. For the class of mechanical problems that we are interested in, for example, the state can be $\mathbf{x} = \{\mathbf{q}^T, \mathbf{v}^T\}^T$, with positions $\mathbf{q}$ and velocities $\mathbf{v}$, and with $\mathbf{u}$ as the set of reaction forces, that must satisfy a VI. References about DVIs can be found in [28, 29]; details about their practical implementations in terms of timestepping schemes can be found, for instance, in [3, 11, 16, 25, 30, 31].

In the following we show that the introduction of set-valued rolling friction does not affect overly the complexity of the original DVI problem; the computational requirements are simply doubled with respect to the case of just sliding friction. Also, spinning friction (also known as *drilling* friction) can be added easily in a similar way.

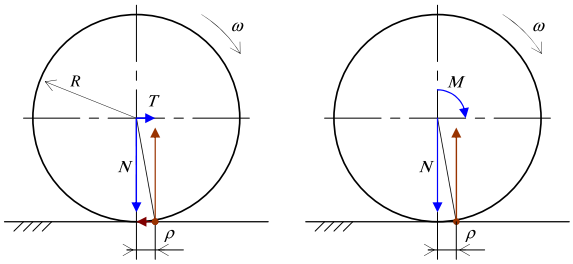

Fig. 1 Rolling friction. Two examples in the two-dimensional case

2 Set-valued rolling friction

In this paper we use set-valued functions to model rolling contact forces between rigid parts. Such model has mathematical similarities with the model for sliding friction in the Amontons-Coulomb theory; similar to the Amontons-Coulomb friction model, the proposed approach requires a small number of parameters to describe the rolling resistance effect.

2.1 Rolling friction phenomena

Although the nonsmooth nature of the sliding friction is evident in nature, because the friction force is abruptly clamped to a maximum value as soon as the objects start sliding, the same sharpness is not evident in the experimental observation of the rolling friction, because rolling friction effect increases smoothly as the rolling speeds increases. In fact, resistance to the rolling motion usually takes place because in the contact area there may be inelastic deformations of elastoplastic materials [24]: the final outcome of this hysteric deformation is that the resultant of all pressures is always placed a bit ahead of the position that it would take if the parts were not rolling.

Many classical textbooks about applied mechanics [8, 26, 37] describe a simplified model for this effect, considering a rolling wheel of radius R and expressing the displacement of the contact force by a single *friction parameter* ρ, which has the dimension of length, or by a dimensionless *coefficient of rolling friction* $f_\rho = \rho/R$. The latter is also meant to allow an easy comparison with sliding friction: Given a normal force N acting on a rolling disc, the horizontal force that is needed to keep it rolling at constant speed is $T = f_\rho N$ (see Fig. 1), similar to the Coloumb sliding friction case, $T = f_s N$, with f_s the coefficient of dry sliding friction. Equivalently, the effect of T on the disc can be replaced with a tractive torque $M = TR$, that is, $M = f_\rho NR$ or

$$M = \rho N. \tag{4}$$

This model is highly nonlinear because it states that the displacement, whose amount is ρ regardless of the speed, has to change direction when the rolling speed changes sign and must go to zero if there is no rolling speed. This tristate model cannot be used practically within a general-purpose numerical simulation framework because scenarios often occur where a sphere or a disc should come to rest over a horizontal surface: since numerical roundoff considerations make impossible that the speed will be exactly null, the speed may actually oscillate around the null value, but even small oscillations will change the sign of the rolling speed and, consequently, also the displacement of the contact force will oscillate over the endpoints of the $-\rho, +\rho$ interval. The final result will be numerically unstable.

For this reason, we express the rolling friction model (4) as the following constraint with inequality

$$\begin{aligned} &\|M\| \le \rho \|N\|, \\ &\|\omega_r\|(\rho N - M) = 0, \quad M\omega_r \le 0, \end{aligned} \tag{5}$$

where ω_r is the rolling angular velocity, which must be opposite to the rolling resistant moment M. The third condition of Eq. (5) can also be written as $\langle M, \omega_r \rangle = -\|M\|\|\omega_r\|$.

Note that for whatever positive or negative rolling speed, this model corresponds to the classical rolling-friction model (4), but for the transition from steady state to moving state it changes the displacement in the $[-\rho, \rho]$ limit. With this improvement, the model can be seen as the counterpart of the Amontons-Coulomb friction model, because both can consider the static case.

2.2 Three dimensional rolling friction model

For extending the two-dimensional rolling friction model to the generic case of contact between shapes in three-dimensional space, we introduce the following assumption.

Assumption A1 The resisting torque is opposite to the relative (rolling) angular velocity of the two bodies, if any.

Although Assumption A1 is always verified for 2D problems, in some three-dimensional problems the resisting torque can be misaligned with respect to the relative angular velocity. See Appendix A for a derivation of the relative angular velocity.

In cases with plane symmetry in the surroundings of the contact (for instance if the area of contact between two rolling bodies is an ellipse aligned to the rolling direction and the material properties are anisothropic), the speed of deformation of the material is symmetric with respect to the plane of rotation, thus resulting in a symmetric pressure distribution at the area of contact, regardless of the type of viscous constitutive law of the material. In this case, Assumption A1 is always verified. This happens, for example, in case of spheres that are in contact, each with anisothropic material, or in spheres that are rolling on flat surfaces.

Another special case that has much relevance in engineering applications is the contact between two surfaces that can be locally approximated as two cylinders with parallel axes. This happens, for instance, in cam followers and in rollers over flat surfaces. In these cases, the torque is aligned to the rolling angular velocity.

In other situations, such as in the case of two generic ellipsoids, the pressure distribution in the area of contact (that is elliptical and not necessarily with one of its main axes aligned to the direction of rolling) generates a normal reaction whose offset with respect to the plane of contact might be not aligned to the direction of rolling, thus resulting in a resisting moment that is not exactly aligned to the vector of the angular velocity.

Given the ith contact, among two bodies A and B, let $\mathbf{n}^i$ be the normal at the contact point, directed toward the exterior of the A body. Let $\mathbf{u}^i$ and $\mathbf{w}^i$ be two vectors in the contact plane such that $\mathbf{n}^i, \mathbf{u}^i, \mathbf{w}^i \in \mathbb{R}^3$ are mutually orthogonal vectors.

The signed gap function Φ^i represents the contact distance. For each contact that is active (that is $\Phi^i(\cdot) = 0$ because bodies are touching), we introduce the contact forces, while inactive contacts ($\Phi^i(\cdot) > 0$) do not enforce any reaction.

The normal contact force is $\mathbf{F}^i_N = \hat{\gamma}^i_n \mathbf{n}^i$, where $\hat{\gamma}^i_n \geq 0$ is the multiplier that represents the modulus of the reaction. Friction force, if any, is represented by the multipliers $\hat{\gamma}^i_u$, and $\hat{\gamma}^i_w$ which lead to the tangential component of the reaction $\mathbf{F}^i_T = \hat{\gamma}^i_u \mathbf{u}^i + \hat{\gamma}^i_w \mathbf{w}^i$.

Because of the inequality $\hat{\gamma}^i_n \geq 0$, the mathematical description of this unilateral model involves the Signorini complementarity problem [30]:

$$\hat{\gamma}^i_n \geq 0 \quad \perp \quad \Phi^i(\cdot) \geq 0. \tag{6}$$

We introduce the rolling friction torque using the multipliers $\hat{\tau}^i_n$, $\hat{\tau}^i_u$, and $\hat{\tau}^i_w$, which correspond to a normal component of the torque $\mathbf{M}^i_N = \hat{\tau}^i_n \mathbf{n}^i$ and two tangential components of the torque $\mathbf{M}^i_T = \hat{\tau}^i_u \mathbf{u}^i + \hat{\tau}^i_w \mathbf{w}^i$.

The normal component of the torque $\mathbf{M}^i_N$ is responsible for friction that reacts to spinning around the vertical axis, while the tangential component $\mathbf{M}^i_T = \hat{\tau}^i_u \mathbf{u}^i + \hat{\tau}^i_w \mathbf{w}^i$ is the effect of the classical rolling friction. The model (5) is extended to the three-dimensional case, the following inequality holds for nonzero rolling velocity.

$$\left\| \mathbf{M}^i_T \right\| \leq \rho \hat{\gamma}^i_n$$

This corresponds to the inequality $\rho^i \hat{\gamma}^i_n \geq \sqrt{\hat{\tau}^{i\,2}_u + \hat{\tau}^{i\,2}_w}$.

The rolling velocity vector ω^i_T is the part of the relative angular velocity vector ω^i_r that lies on the contact plane, that is, $\omega^i_T = \omega^i_r - \mathbf{n}^i \langle \omega^i_r, \mathbf{n}^i \rangle$; the condition that requires ω^i_T to be aligned and opposite to $\mathbf{M}^i_T$ can be expressed as

$$\langle \mathbf{M}^i_T, \omega^i_T \rangle = -\left\| \mathbf{M}^i_T \right\| \left\| \omega^i_T \right\|.$$

Therefore, the full rolling friction model for a contact with rolling friction parameter ρ^i is mathematically equivalent to the following constraints:

$$\begin{cases} \rho^i \hat{\gamma}^i_n \geq \sqrt{\hat{\tau}^{i\,2}_u + \hat{\tau}^{i\,2}_w} \\ \|\omega^i_T\| (\rho^i \hat{\gamma}^i_n - \sqrt{\hat{\tau}^{i\,2}_u + \hat{\tau}^{i\,2}_w}) = 0, \\ \langle \mathbf{M}^i_T, \omega^i_T \rangle = -\|\mathbf{M}^i_T\| \, \|\omega^i_T\|. \end{cases} \tag{7}$$

Additionally, one can introduce the spinning friction, represented by a parameter σ_i, giving

$$\begin{cases} \sigma^i \hat{\gamma}^i_n \geq \hat{\tau}^i_n \\ \|\omega^i_N\| (\sigma^i \hat{\gamma}^i_n - \hat{\tau}^i_n) = 0, \\ \langle \mathbf{M}^i_N, \omega^i_N \rangle = -\|\mathbf{M}^i_N\| \, \|\omega^i_N\|. \end{cases} \tag{8}$$

Rolling contacts can be either sliding or not sliding. In the former case there could be also tangential forces caused by dynamical friction; in the latter case there could be forces caused by sticking, the consequence of static friction. Therefore we must introduce the Amontons-Coulomb friction model to take care of the tangential forces, either sliding or static.

Within the classic theory of dry friction, the friction coefficient μ^i limits the ratio between the normal and the tangential force, and the tangential force must have a direction that is opposite to $\mathbf{v}_T^i$, the tangential component of the relative velocity $\mathbf{v}_r^i$, if any, thus requiring

$$\begin{cases} \mu^i \hat{\gamma}_n^i \geq \sqrt{\hat{\gamma}_u^{i\,2} + \hat{\gamma}_w^{i\,2}} \\ \|\mathbf{v}_T^i\|(\mu^i \hat{\gamma}_n^i - \sqrt{\hat{\gamma}_u^{i\,2} + \hat{\gamma}_w^{i\,2}}) = 0, \\ \langle \mathbf{F}_T^i, \mathbf{v}_T^i \rangle = -\|\mathbf{F}_T^i\| \, \|\mathbf{v}_T^i\|. \end{cases} \tag{9}$$

3 The complete DVI model

The system state is defined by the vector of generalized coordinates $\mathbf{q} \in \mathbb{R}^{m_q}$ and the vector of generalized speeds $\mathbf{v} \in \mathbb{R}^{m_v}$. It might happen that $m_q > m_v$ because rotations of rigid bodies in three-dimensional space are represented with unimodular quaternions $\epsilon \in \mathbb{H}_1$ to avoid singularities in the parametrization of $\mathsf{SO}(\mathbb{R}, 3)$; anyway it is straightforward to define a (linear) map $\dot{\mathbf{q}} = \Gamma(\mathbf{q})\mathbf{v}$ if $\dot{\mathbf{q}}$ is needed.

We also introduce generalized force fields $\mathbf{f}_e(\mathbf{q}, \mathbf{v}, t)$ and gyroscopic forces $\mathbf{f}_c(\mathbf{q}, \mathbf{v})$ giving a total force field $\mathbf{f}_t(\mathbf{q}, \mathbf{v}, t) \in \mathbb{R}^{m_v}$.

The inertial properties of the system are represented by the mass matrix $M(\mathbf{q}) \in \mathbb{R}^{m_v \times m_v}$, assumed positive definite, usually block-diagonal in the case of rigid bodies only.

Bilateral constraints are introduced through a set $\mathscr{B}$ of scalar constraint equations, assumed differentiable everywhere:

$$\Psi^i(\mathbf{q}, t) = 0, \quad i \in \mathscr{B}. \tag{10}$$

We introduce $\nabla_q \Psi^i = [\partial \Psi^i / \partial \mathbf{q}]^T$ and $\nabla \Psi^{i\,T} = \nabla_q \Psi^{i\,T} \Gamma(\mathbf{q})$, to express the constraint (10) at the velocity level after differentiation:

$$\frac{d\Psi^i(\mathbf{q}, t)}{dt} = \nabla \Psi^{i\,T} \mathbf{v} + \frac{\partial \Psi^i}{\partial t} = 0, \quad i \in \mathscr{B}. \tag{11}$$

Frictional unilateral contacts define a set $\mathscr{A}$. For each contact $i \in \mathscr{A}$, we introduce the tangent space generators $\mathbf{D}_{\gamma_u}^i, \mathbf{D}_{\gamma_w}^i, \mathbf{D}_{\tau_u}^i, \mathbf{D}_{\tau_w}^i, \mathbf{D}_{\tau_n}^i \in \mathbb{R}^{m_v}$; for details about their formulation, see Appendix B.

Another way to write (7), (8), and (9) is to use the maximum dissipation principle, thus leading respectively to the following constraints on the dynamic equilibrium

$$\begin{aligned} &(\hat{\tau}_u^i, \hat{\tau}_w^i) = \operatorname{argmin} \mathbf{v}^T (\mathbf{D}_{\tau_u}^i \hat{\tau}_u^i + \mathbf{D}_{\tau_w}^i \hat{\tau}_w^i) \\ &\text{s.t. } (\hat{\tau}_u^i, \hat{\tau}_w^i) \in \mathscr{Z}_r^i \\ &\mathscr{Z}_r^i \triangleq \{(\hat{\tau}_u^i, \hat{\tau}_w^i) \big| \sqrt{\hat{\tau}_u^{i\,2} + \hat{\tau}_w^{i\,2}} \leq \rho^i \hat{\gamma}_n^i\} \end{aligned} \tag{12}$$

$$\begin{aligned} &(\hat{\tau}_n^i) = \operatorname{argmin} \mathbf{v}^T (\mathbf{D}_{\tau_n}^i \hat{\tau}_n^i) \\ &\text{s.t. } \hat{\tau}_n^i \in \mathscr{Z}_s^i \\ &\mathscr{Z}_s^i \triangleq \{\hat{\tau}_n^i \big| |\hat{\tau}_n^i| \leq \sigma^i \hat{\gamma}_n^i\} \end{aligned} \tag{13}$$

$$\begin{aligned} &(\hat{\gamma}_u^i, \hat{\gamma}_w^i) = \operatorname{argmin} \mathbf{v}^T (\mathbf{D}_{\gamma_u}^i \hat{\gamma}_u^i + \mathbf{D}_{\gamma_w}^i \hat{\gamma}_w^i) \\ &\text{s.t. } (\hat{\gamma}_u^i, \hat{\gamma}_w^i) \in \mathscr{Z}_f^i \\ &\mathscr{Z}_f^i \triangleq \{(\hat{\gamma}_u^i, \hat{\gamma}_w^i) \big| \sqrt{\hat{\gamma}_u^{i\,2} + \hat{\gamma}_w^{i\,2}} \leq \mu^i \hat{\gamma}_n^i\} \end{aligned} \tag{14}$$

Here, $\mathscr{Z}_r^i$, $\mathscr{Z}_s^i$, and, respectively, $\mathscr{Z}_f^i$, are the rolling, spinning, and respectively, friction cone sections. Alternatively, for the case where the tangential rolling torques is not zero, we derive the Fritz John optimality conditions for the nonlinear program (12),

$$\begin{aligned} s = \nabla_{\tau_u, \tau_w} \mathbf{v}^T (\mathbf{D}_{\tau_u} \hat{\tau}_u^i + \mathbf{D}_{\tau_w} \hat{\tau}_w^i) \\ - \lambda_\omega \nabla_{\tau_u, \tau_w} (\rho \hat{\gamma}_n^i - \sqrt{\hat{\tau}_u^{i\,2} + \hat{\tau}_w^{i\,2}}) = 0 \end{aligned} \tag{15}$$

$$\rho^i \hat{\gamma}_n^i - \sqrt{\hat{\tau}_u^{i\,2} + \hat{\tau}_w^{i\,2}} \geq 0, \quad \perp \quad \lambda_\omega \geq 0. \tag{16}$$

When the tangential rolling torques are zero, the functions defining $\mathscr{Z}_r^i$, are not differentiable so the above derivation does not hold. However, for purposes of exposing our formulation we will assume the opposite the case, and in later developments (40) we remove this assumptions by using cone polar formalisms.

The same derivation can be performed for (12) and (13); hence the complete model, including inertial effects, force fields, bilateral constraints, unilateral contacts with friction, rolling friction, and spinning friction, is the following differential variational inequality

$$
\begin{aligned}
&\dot{\mathbf{q}} = \Gamma(\mathbf{q})\mathbf{v}\\
&M(\mathbf{q})\frac{d\mathbf{v}}{dt} = \sum_{i\in\mathscr{A}}\begin{pmatrix}\hat{\gamma}_n^i\mathbf{D}_{\gamma_n}^i + \hat{\gamma}_u^i\mathbf{D}_{\gamma_u}^i + \hat{\gamma}_w^i\mathbf{D}_{\gamma_w}^i\\ + \hat{\tau}_n^i\mathbf{D}_{\tau_n}^i + \hat{\tau}_u^i\mathbf{D}_{\tau_u}^i + \hat{\tau}_w^i\mathbf{D}_{\tau_w}^i\end{pmatrix}\\
&\qquad + \sum_{i\in\mathscr{B}}\hat{\gamma}_{\mathscr{B}}^i\nabla\Psi^i + \mathbf{f}_t(t,\mathbf{q},\mathbf{v})\\
&i\in\mathscr{B}:\quad \Psi^i(\mathbf{q},t) = 0\\
&i\in\mathscr{A}:\quad \hat{\gamma}_n^i \geq 0 \quad\perp\quad \Phi^i(\mathbf{q}) \geq 0\\
&\nabla_{\tau_u,\tau_w}\mathbf{v}^T\left(\mathbf{D}_{\tau_u}\hat{\tau}_u^i + \mathbf{D}_{\tau_w}\hat{\tau}_w^i\right)\\
&\quad - \lambda_\omega^i\nabla_{\tau_u,\tau_w}\left(\rho^i\hat{\gamma}_n^i - \sqrt{\hat{\tau}_u^{i\,2} + \hat{\tau}_w^{i\,2}}\right) = 0\\
&\rho^i\hat{\gamma}_n^i - \sqrt{\hat{\tau}_u^{i\,2} + \hat{\tau}_w^{i\,2}} \geq 0, \quad\perp\quad \lambda_\omega^i \geq 0\\
&\nabla_{\gamma_u,\gamma_w}\mathbf{v}^T\left(\mathbf{D}_{\gamma_u}\hat{\gamma}_u^i + \mathbf{D}_{\gamma_w}\hat{\gamma}_w^i\right)\\
&\quad - \lambda_v^i\nabla_{\gamma_u,\gamma_w}\left(\mu^i\hat{\gamma}_n^i - \sqrt{\hat{\gamma}_u^{i\,2} + \hat{\gamma}_w^{i\,2}}\right) = 0\\
&\mu^i\hat{\gamma}_n^i - \sqrt{\hat{\gamma}_u^{i\,2} + \hat{\gamma}_w^{i\,2}} \geq 0, \quad\perp\quad \lambda_v^i \geq 0\\
&\nabla_{\tau_n}\mathbf{v}^T\left(\mathbf{D}_{\tau_n}\hat{\tau}_n^i\right) - \lambda_\tau^i\nabla_{\tau_n}\left(\sigma^i\hat{\gamma}_n^i - \left|\hat{\tau}_n^i\right|\right) = 0\\
&\mu^i\hat{\gamma}_n^i - \left|\hat{\tau}_n^i\right| \geq 0, \quad\perp\quad \lambda_\tau^i \geq 0
\end{aligned}
\tag{17}
$$

The former DVI can be discretized in time. Using a time step h, posing $\gamma = h\hat{\gamma}$, and adopting the exponential map $\Lambda(\cdot)$ described in [33] to allow direct integration on the Lie group, we have the following problem.

$$
\begin{aligned}
&\mathbf{q}^{(l+1)} = \Lambda\left(\mathbf{q}^{(l)}, \mathbf{v}^{(l+1)}, h\right)\\
&M\mathbf{v}^{(l+1)} = \sum_{i\in\mathscr{A}}\begin{pmatrix}\gamma_n^i\mathbf{D}_{\gamma_n}^i + \gamma_u^i\mathbf{D}_{\gamma_u}^i + \gamma_w^i\mathbf{D}_{\gamma_w}^i\\ + \tau_n^i\mathbf{D}_{\tau_n}^i + \tau_u^i\mathbf{D}_{\tau_u}^i + \tau_w^i\mathbf{D}_{\tau_w}^i\end{pmatrix}\\
&\qquad + \sum_{i\in\mathscr{B}}\gamma_{\mathscr{B}}^i\nabla\Psi^i + h\mathbf{f}_t(t,\mathbf{q},\mathbf{v}) + M\mathbf{v}^{(l)}\\
&i\in\mathscr{B}:\quad \frac{1}{h}\Psi^i\left(\mathbf{q}^{(l)}\right) + \nabla\Psi^{i\,T}\mathbf{v}^{(l+1)} + \frac{\partial\Psi^i}{\partial t} = 0\\
&i\in\mathscr{A}:\quad \gamma_n^i \geq 0 \quad\perp\quad \frac{1}{h}\Phi^i\left(\mathbf{q}^{(l)}\right) + \nabla\Phi^{i\,T}\mathbf{v}^{(l+1)} \geq 0\\
&\nabla_{\tau_u,\tau_w}\mathbf{v}^T\left(\mathbf{D}_{\tau_u}\tau_u^i + \mathbf{D}_{\tau_w}\tau_w^i\right)\\
&\quad - \lambda_\omega^i\nabla_{\tau_u,\tau_w}\left(\rho^i\gamma_n^i - \sqrt{\tau_u^{i\,2} + \tau_w^{i\,2}}\right) = 0\\
&\rho^i\gamma_n^i - \sqrt{\tau_u^{i\,2} + \tau_w^{i\,2}} \geq 0, \quad\perp\quad \lambda_\omega^i \geq 0\\
&\nabla_{\gamma_u,\gamma_w}\mathbf{v}^T\left(\mathbf{D}_{\gamma_u}\gamma_u^i + \mathbf{D}_{\gamma_w}\gamma_w^i\right)\\
&\quad - \lambda_v^i\nabla_{\gamma_u,\gamma_w}\left(\mu^i\gamma_n^i - \sqrt{\gamma_u^{i\,2} + \gamma_w^{i\,2}}\right) = 0\\
&\mu^i\gamma_n^i - \sqrt{\gamma_u^{i\,2} + \gamma_w^{i\,2}} \geq 0, \quad\perp\quad \lambda_v^i \geq 0\\
&\nabla_{\tau_n}\mathbf{v}^T\left(\mathbf{D}_{\tau_n}\tau_n^i\right) - \lambda_\tau^i\nabla_{\tau_n}\left(\sigma^i\gamma_n^i - \left|\tau_n^i\right|\right) = 0\\
&\mu^i\gamma_n^i - \left|\tau_n^i\right| \geq 0, \quad\perp\quad \lambda_\tau^i \geq 0
\end{aligned}
\tag{18}
$$

This is a mixed nonlinear complementarity problem, whose solution is not guaranteed to exist. Indeed, most existence results require monotonicity of the mapping defining the complementarity problem. In turn, this implies convexity of the solution set of the nonlinear complementarity problem [10]. Unfortunately, not even the weaker condition of the convexity of the solution set can be guaranteed. Indeed, it has been already shown that, for the subcase of a linear complementarity problem (LCP) corresponding to a simple pyramidal frictional model, the solution set may be nonconvex [1]. This situation can occur only if the mapping of the LCP is nonmonotone, which in the linear case implies that its matrix is not positive semi-definite.

4 Casting to a convex solvable problem

In the following we show that, under mild conditions, the original problem can be relaxed as a monotone cone complementarity problem (CCP) that guarantees existence of the solution and convexity of the solution set. The problem is equivalent to a convex optimization problem, a feasible quadratic programming problem (QP) with conical constraints. We note that while one cannot guarantee uniqueness of the solution of the CCP (indeed, lack of uniqueness of the forces is known to appear even in frictionless cases that are linear and convex problems), one can guarantee under some mild assumptions uniqueness of the velocity solution [2]. The latter condition is sufficient for a predictive time-stepping scheme to exist.

Assume that for each contact $i \in \mathscr{A}$ we define the vector of wrench reactions

$$
\boldsymbol{\gamma}_{\mathscr{A}}^i = \left\{\gamma_n^i, \gamma_u^i, \gamma_w^i, \tau_n^i, \tau_u^i, \tau_w^i\right\}^T
$$

and the corresponding twist of local linear and angular velocities in the contact point, plus the stabilization term $\frac{1}{h}\Phi(\mathbf{q})$, that is,

$$
\begin{aligned}
\mathbf{u}_{\mathscr{A}}^i = \Bigg\{&\nabla\Phi^{i,T}\mathbf{v} + \frac{1}{h}\Phi^i,\\
&\mathbf{D}_{\gamma_u}^{i,T}\mathbf{v}, \mathbf{D}_{\gamma_v}^{i,T}\mathbf{v}, \mathbf{D}_{\tau_n}^{i,T}\mathbf{v}, \mathbf{D}_{\tau_u}^{i,T}\mathbf{v}, \mathbf{D}_{\tau_w}^{i,T}\mathbf{v}\Bigg\}^T.
\end{aligned}
$$

In the rest of this section we omit the i indexes for compactness.

We differentiate the Fritz John optimality conditions (15) obtaining, for the rolling friction part,

$$\mathbf{v}^T\mathbf{D}_{\tau_u} = \lambda_\omega \tau_u \frac{1}{\sqrt{\tau_u^2+\tau_w^2}} \tag{19}$$

$$\mathbf{v}^T\mathbf{D}_{\tau_w} = \lambda_\omega \tau_w \frac{1}{\sqrt{\tau_u^2+\tau_w^2}}. \tag{20}$$

Then,

$$\lambda_\omega = \sqrt{\left(\mathbf{D}_{\tau_u}^T\mathbf{v}\right)^2 + \left(\mathbf{D}_{\tau_w}^T\mathbf{v}\right)^2}. \tag{21}$$

Note that for the sliding friction one gets similar results, $\mathbf{v}^T\mathbf{D}_{\gamma_u} = \lambda_v \gamma_u 1/\sqrt{\gamma_u^2+\gamma_w^2}$ and $\mathbf{v}^T\mathbf{D}_{\gamma_w} = \lambda_v \gamma_w 1/\sqrt{\gamma_u^2+\gamma_w^2}$, with $\lambda_v = \sqrt{(\mathbf{D}_{\gamma_u}^T\mathbf{v})^2 + (\mathbf{D}_{\gamma_w}^T\mathbf{v})^2}$; whereas for the spinning friction $\mathbf{v}^T\mathbf{D}_{\tau_n} = \lambda_\tau \tau_n/|\tau_n|$, with $\lambda_\tau = \mathbf{D}_{\tau_n}^T\mathbf{v}|\tau_n|/\tau_n$.

Now, the inner product for the complementarity in the optimality condition requires that

$$\lambda_\omega\left(\rho\gamma_n - \sqrt{\tau_u^2+\tau_w^2}\right) = 0$$

$$\lambda_\omega\sqrt{\tau_u^2+\tau_w^2} = \lambda_\omega \rho\gamma_n. \tag{22}$$

Similarly one can derive, for sliding and spinning,

$$\lambda_v\sqrt{\gamma_u^2+\gamma_w^2} = \lambda_v \mu\gamma_n \tag{23}$$

$$\lambda_\tau|\tau_n| = \lambda_\tau \sigma\gamma_n. \tag{24}$$

If we assume the orthogonality of $\boldsymbol{\gamma}_{\mathscr{A}}$ and $\mathbf{u}_{\mathscr{A}}$, we have

$$\begin{aligned}&\langle \boldsymbol{\gamma}_{\mathscr{A}}, \mathbf{u}_{\mathscr{A}}\rangle \\ &= \gamma_n\left(\nabla\Phi^T\mathbf{v} + \frac{1}{h}\Phi\right) + \gamma_u\mathbf{D}_{\gamma_u}^T\mathbf{v} + \gamma_w\mathbf{D}_{\gamma_w}^T\mathbf{v} \\ &\quad + \tau_n\mathbf{D}_{\tau_n}^T\mathbf{v} + \tau_u\mathbf{D}_{\tau_u}^T\mathbf{v} + \tau_w\mathbf{D}_{\tau_w}^T\mathbf{v} = 0.\end{aligned} \tag{25}$$

By adding a relaxation term A_r for the normal velocity, the orthogonality condition for unilateral contact in the discretized DVI (18) becomes

$$\gamma_n \geq 0 \quad \perp \quad \frac{1}{h}\Phi + \nabla\Phi^T\mathbf{v} + A_r \geq 0.$$

Since complementarity implies nullity of inner product, we then have

$$\gamma_n\left(\frac{1}{h}\Phi + \nabla\Phi^T\mathbf{v}\right) = -\gamma_n A_r \geq 0. \tag{26}$$

By exploiting (26), (19), (20), and so forth, recalling that $x^2/|x| = |x|$, with simple algebraic manipulation, we rewrite (25) as

$$\begin{aligned}\langle \boldsymbol{\gamma}_{\mathscr{A}}, \mathbf{u}_{\mathscr{A}}\rangle &= -\gamma_n A_r + \frac{\gamma_u^2\lambda_v}{\sqrt{\gamma_u^2+\gamma_w^2}} + \frac{\gamma_w^2\lambda_v}{\sqrt{\gamma_u^2+\gamma_w^2}} \\ &\quad + \frac{\tau_u^2\lambda_\omega}{\sqrt{\tau_u^2+\tau_w^2}} + \frac{\tau_w^2\lambda_\omega}{\sqrt{\tau_u^2+\tau_w^2}} \\ &\quad + \frac{\tau_n^2\lambda_\tau}{|\tau_n|} = 0 \\ \gamma_n A_r &= \lambda_v\sqrt{\gamma_u^2+\gamma_w^2} + \lambda_\omega\sqrt{\tau_u^2+\tau_w^2} + \lambda_\tau|\tau_n|.\end{aligned} \tag{27}$$

By substituting (22), (23), and (24) in (27), and by simplifying γ_n, we have that the relaxation term that allows orthogonality of $\boldsymbol{\gamma}$'s and $\mathbf{u}$'s is, for the ith contact $i \in \mathscr{A}$,

$$\begin{aligned}A_r^i &= \mu^i\sqrt{\left(\mathbf{D}_{\gamma_u}^i\mathbf{v}\right)^2 + \left(\mathbf{D}_{\gamma_w}^i\mathbf{v}\right)^2} \\ &\quad + \rho^i\sqrt{\left(\mathbf{D}_{\tau_u}^i\mathbf{v}\right)^2 + \left(\mathbf{D}_{\tau_w}^i\mathbf{v}\right)^2} + \sigma^i\left|\mathbf{D}_{\tau_n}^i\mathbf{v}\right|.\end{aligned} \tag{28}$$

The introduction of the A_r^i term has the drawback of modifying the contact constraint; from a practical point of view there is the side effect that the gap between objects increase with the sliding speed instead of remaining equal to zero; this effect has been discussed in [33] for the simple case of sliding friction. Now, introducing also rolling and spinning friction, one can see from Eq. (28) that rotational motion increases the gap (the $\sqrt{(\mathbf{D}_{\tau_u}^i\mathbf{v})^2 + (\mathbf{D}_{\tau_w}^i\mathbf{v})^2}$ term is also the norm of ω_T^i), and that the increase is always directed outward. This is the price for having convexified the original problem. In many situations this can be acceptable, indeed one can demonstrate that at steady state, i.e. when $\nabla\Phi^{i\,T}\mathbf{v} = 0$ in (26), the separation gap Φ^i decreases with low speeds, low μ, ρ, σ, and small timesteps.

We note that if we plan to use rolling friction to simulate granular material, the above mentioned side effect of the relaxation leads to a dilatancy effect that really happens in physical world.

Aiming at a generic compact notation, we now introduce $\mathbf{b}_{\mathscr{A}}^i \in \mathbb{R}^6 = \{\frac{1}{h}\Phi^i, 0, 0, 0, 0, 0\}^T$, and we build the following aggregate vectors $\mathbf{b}_{\mathscr{A}} \in \mathbb{R}^{6n_{\mathscr{A}}}$, $\boldsymbol{\gamma}_{\mathscr{A}} \in \mathbb{R}^{6n_{\mathscr{A}}}$, $\mathbf{u}_{\mathscr{A}} \in \mathbb{R}^{6n_{\mathscr{A}}}$:

$$\begin{aligned}\mathbf{b}_{\mathscr{A}} &= \left\{\mathbf{b}_{\mathscr{A}}^{1,T}, \mathbf{b}_{\mathscr{A}}^{2,T}, \ldots, \mathbf{b}_{\mathscr{A}}^{n_{\mathscr{A}},T}\right\}^T, \\ \boldsymbol{\gamma}_{\mathscr{A}} &= \left\{\boldsymbol{\gamma}_{\mathscr{A}}^{1,T}, \boldsymbol{\gamma}_{\mathscr{A}}^{2,T}, \ldots, \boldsymbol{\gamma}_{\mathscr{A}}^{n_{\mathscr{A}},T}\right\}^T, \\ \mathbf{u}_{\mathscr{A}} &= \left\{\mathbf{u}_{\mathscr{A}}^{1^T}, \mathbf{u}_{\mathscr{A}}^{2^T}, \ldots, \mathbf{u}_{\mathscr{A}}^{n_{\mathscr{A}},T}\right\}^T.\end{aligned} \tag{29}$$

For each contact we can define a matrix with six columns,

$$D^i = \left[\nabla\Phi^i \middle| \mathbf{D}^i_{\gamma_u} \middle| \mathbf{D}^i_{\gamma_w} \middle| \mathbf{D}^i_{\tau_n} \middle| \mathbf{D}^i_{\tau_u} \middle| \mathbf{D}^i_{\tau_w}\right], \tag{30}$$

and a six-dimensional cone that defines the set of admissible reactions in the sliding, rolling, spinning friction contact:

$$\mathscr{Z}^i = \left\{ \boldsymbol{\gamma} \in \mathbb{R}^6 \;\middle|\; \left\{ \begin{array}{l} \mu^i \gamma_n \geq \sqrt{\gamma_u^2 + \gamma_w^2}, \\ \rho^i \gamma_n \geq \sqrt{\tau_u^2 + \tau_w^2}, \\ \sigma^i \gamma_n \geq |\tau_n| \end{array} \right\} \right\}. \tag{31}$$

Similarly, for bilateral constraints, we have $\mathbf{b}_{\mathscr{B}} \in \mathbb{R}^{n_{\mathscr{B}}}$, and $\boldsymbol{\gamma}_{\mathscr{B}} \in \mathbb{R}^{n_{\mathscr{B}}}$:

$$\begin{aligned} \mathbf{b}_{\mathscr{B}} &= \left\{ \frac{1}{h}\Psi^1 + \frac{\partial \Psi^1}{\partial t}, \frac{1}{h}\Psi^2 + \frac{\partial \Psi^2}{\partial t}, \ldots, \right. \\ &\qquad \left. \frac{1}{h}\Psi^{n_{\mathscr{B}}} + \frac{\partial \Psi^{n_{\mathscr{B}}}}{\partial t} \right\}^T \\ \boldsymbol{\gamma}_{\mathscr{B}} &= \left\{ \gamma^1_{\mathscr{B}}, \gamma^2_{\mathscr{B}}, \ldots, \gamma^{n_{\mathscr{B}}}_{\mathscr{B}} \right\}^T \\ \mathbf{u}_{\mathscr{B}} &= \left\{ u^1_{\mathscr{B}}, u^2_{\mathscr{B}}, \ldots, u^{n_{\mathscr{B}}}_{\mathscr{B}} \right\}^T. \end{aligned} \tag{32}$$

The complete aggregate vectors and matrices of the entire system are

$$\begin{aligned} &\boldsymbol{\gamma}_{\mathscr{S}} = \left\{ \boldsymbol{\gamma}^T_{\mathscr{A}}, \boldsymbol{\gamma}^T_{\mathscr{B}} \right\}^T, \qquad \mathbf{u}_{\mathscr{S}} = \left\{ \mathbf{u}^T_{\mathscr{A}}, \mathbf{u}^T_{\mathscr{B}} \right\}^T, \\ &\mathbf{b}_{\mathscr{S}} = \left\{ \mathbf{b}^T_{\mathscr{A}}, \mathbf{b}^T_{\mathscr{B}} \right\}^T, \end{aligned} \tag{33}$$

$$D_{\mathscr{S}} = \left[D^{i_1} \middle| D^{i_2} \middle| \ldots \middle| D^{i_{n_{\mathscr{A}}}} \middle| \nabla\Psi^1 \middle| \nabla\Psi^2 \middle| \ldots \middle| \nabla\Psi^{n_{\mathscr{B}}} \right]. \tag{34}$$

Moreover we define the product of all the sliding and rolling friction cones and the possible values of reactions in bilateral constraints as

$$\Upsilon = \left(\times_{i \in \mathscr{A}} \mathscr{Z}^i \right) \times \left(\times_{i \in \mathscr{B}} \mathbb{R} \right) \tag{35}$$

and its polar (note that $\mathbb{R}^\circ = \{0\}$):

$$\Upsilon^\circ = \left(\times_{i \in \mathscr{A}} \mathscr{Z}^{i^\circ} \right) \times \left(\times_{i \in \mathscr{B}} \{0\} \right). \tag{36}$$

Now we proceed with a simplification by introducing

$$\tilde{\mathbf{k}}^{(l)} = M^{(l)} \mathbf{v}^{(l)} + h\mathbf{f}_t\left(t^{(l)}, \mathbf{q}^{(l)}, \mathbf{v}^{(l)}\right). \tag{37}$$

From the relaxed version of (18) one can see that $\mathbf{u}_{\mathscr{S}} = N\boldsymbol{\gamma}_{\mathscr{S}} + \mathbf{r}$, where

$$N = D^T_{\mathscr{S}} {M^{(l)}}^{-1} D_{\mathscr{S}} \tag{38}$$

$$\mathbf{r} = D^T_{\mathscr{S}} {M^{(l)}}^{-1} \tilde{\mathbf{k}} + \mathbf{b}_{\mathscr{S}}. \tag{39}$$

The entire system is described by the following CCP:

$$(N\boldsymbol{\gamma}_{\mathscr{S}} + \mathbf{r}) \in -\Upsilon^\circ \quad \perp \quad \boldsymbol{\gamma}_{\mathscr{S}} \in \Upsilon. \tag{40}$$

The CCP (40) is also equivalent to a variational inequality as expressed in the VI of Eq. (1), namely

$$\boldsymbol{\gamma}_{\mathscr{S}} \in \Upsilon: \quad \langle N\boldsymbol{\gamma}_{\mathscr{S}} + \mathbf{r}, \mathbf{y} - \boldsymbol{\gamma}_{\mathscr{S}} \rangle \geq 0 \quad \forall \mathbf{y} \in \Upsilon. \tag{41}$$

Several theoretical results for (41) can be obtained by noting that it is related to the following optimization problem.

$$\min_{\boldsymbol{\gamma}_{\mathscr{S}} \in \Upsilon} \frac{1}{2} \boldsymbol{\gamma}^T_{\mathscr{S}} N \boldsymbol{\gamma}_{\mathscr{S}} + \mathbf{r}^T \boldsymbol{\gamma}_{\mathscr{S}} \tag{42}$$

Theorem 1 *Consider the variational inequality* (41) *and the optimization problem* (42). *Then the following statements hold*:

(i) *If* (42) *has a solution, then that solution satisfies the variational inequality* (41). *Conversely, any solution of* (41) *is a solution of* (42).
(ii) *Either* (42) *has a solution, or there exists* $\tilde{\boldsymbol{\gamma}}_{\mathscr{S}}$ *such that* $D_{\mathscr{S}}\tilde{\boldsymbol{\gamma}}_{\mathscr{S}} = 0$.
(iii) *If* (42) *has two solutions* $\boldsymbol{\gamma}^1_{\mathscr{S}}$ *and* $\boldsymbol{\gamma}^2_{\mathscr{S}}$, *then they must satisfy* $D_{\mathscr{S}}\boldsymbol{\gamma}^1_{\mathscr{S}} = D_{\mathscr{S}}\boldsymbol{\gamma}^2_{\mathscr{S}}$.

Proof If (42) has a solution, then since the constraints Υ are convex, they must satisfy the Kuhn-Tucker conditions, which are simply the statement of (41). Conversely, because of the convexity of the cone Υ and of the objective function, any Kuhn-Tucker solution is a global solution, which proves Part (i).

For Part (ii), if the objective function of (42) is bounded below over the convex set Υ, it follows that the problem (42) must have a solution. Assume that this is not the case, that is, that there exists a sequence $\boldsymbol{\gamma}^n_{\mathscr{S}} \in \Upsilon$, $n = 0, 1, 2, \ldots$ such that $\frac{1}{2}\boldsymbol{\gamma}^{nT}_{\mathscr{S}} N \boldsymbol{\gamma}^n_{\mathscr{S}} + \mathbf{r}^T \boldsymbol{\gamma}^n_{\mathscr{S}} \to -\infty$. Because of the continuity of the objective function, such a sequence must satisfy $\|\boldsymbol{\gamma}^n_{\mathscr{S}}\| \to \infty$ (otherwise the objective function would stay bounded). In particular, this implies that

$$\frac{1}{2}\boldsymbol{\gamma}^{nT}_{\mathscr{S}} N \boldsymbol{\gamma}^n_{\mathscr{S}} + \mathbf{r}^T \boldsymbol{\gamma}^n_{\mathscr{S}} \leq 0; \quad \forall n \geq n_0 \tag{43}$$

for some integer n_0. Define the scaled vector $\tilde{\boldsymbol{\gamma}}^n_{\mathscr{S}} = \frac{\boldsymbol{\gamma}^n_{\mathscr{S}}}{\|\boldsymbol{\gamma}^n_{\mathscr{S}}\|} \in \Upsilon$. Since the scaled vector sequence is in the unit ball, which is compact, it must have a limit point, which we denote by $\tilde{\boldsymbol{\gamma}}^*_{\mathscr{S}}$. We assume, without loss of generality, that the entire sequence converges to this point. Dividing (43) by $\|\boldsymbol{\gamma}^n_{\mathscr{S}}\|^2$ and taking the limit, we have that the second term goes to 0, and we obtain

$$\frac{1}{2}\boldsymbol{\gamma}^{*T}_{\mathscr{S}} N \boldsymbol{\gamma}^{*}_{\mathscr{S}}. \leq 0$$

Using the expression for N, (38), and the fact that the mass matrix is positive definite, this implies that $D_{\mathscr{S}}\tilde{\boldsymbol{\gamma}}^{*}_{\mathscr{S}} = 0$, which in turn proves Part (ii) of the theorem.

For Part (iii), assume that there are two solutions of the optimization problem (42), $\boldsymbol{\gamma}^{1}_{\mathscr{S}}$ and $\boldsymbol{\gamma}^{2}_{\mathscr{S}}$. Then, by the fact that the objective function is convex and that the constraints are convex, for any $t \in [0, 1]$, we have that $\boldsymbol{\gamma}^{1}_{\mathscr{S}} + t(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1})$ is also a solution of (42). That is, the objective function value

$$\begin{aligned}&\left(\boldsymbol{\gamma}^{1}_{\mathscr{S}} + t\left(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1}\right)\right)^{T} N\left(\boldsymbol{\gamma}^{1}_{\mathscr{S}} + t\left(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1}\right)\right)\\ &\quad + \mathbf{r}^{T}\left(\boldsymbol{\gamma}^{1}_{\mathscr{S}} + t\left(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1}\right)\right)\end{aligned}$$

is constant in t. This function is a quadratic, and this can occur only if the coefficient in t^2 is 0, that is,

$$\left(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1}\right)^{T} N\left(\boldsymbol{\gamma}^{2}_{\mathscr{S}} - \boldsymbol{\gamma}^{1}\right) = 0.$$

Again using the expression for N, (38), and the fact that the mass matrix is positive definite, this implies that $D_{\mathscr{S}}(\tilde{\boldsymbol{\gamma}}^{2}_{\mathscr{S}} - \tilde{\boldsymbol{\gamma}}^{1}_{\mathscr{S}}) = 0$, which in turn proves Part (iii) of the theorem. The proof of the theorem is complete. □

After the dual variables $\boldsymbol{\gamma}_{\mathscr{S}}$ have been obtained from the VI (41), one can easily compute also the primal variables with the affine mapping:

$$\mathbf{v}^{(l+1)} = {M^{(l)}}^{-1} D_{\mathscr{S}}\boldsymbol{\gamma}_{\mathscr{S}} + {M^{(l)}}^{-1}\tilde{\mathbf{k}}. \tag{44}$$

In what sense the problem has a unique solution now revolves around the results of Theorem 1. The key assumption that we make is that the friction cone of our problem is *pointed*, that is,

$$\boldsymbol{\gamma}_{\mathscr{S}} \in \Upsilon, \quad D_{\mathscr{S}}\boldsymbol{\gamma}_{\mathscr{S}} = 0 \quad \Rightarrow \quad \boldsymbol{\gamma}_{\mathscr{S}} = 0. \tag{45}$$

Note that Υ is only a set of multipliers (that is, the contact constraint forces in the coordinates attached to the contact, not the bodies) that is mapped into friction forces (and the friction cone, that is, in general coordinate) by means of the mapping $D_{\mathscr{S}}\boldsymbol{\gamma}_{\mathscr{S}}$, and thus (45) is a statement about friction forces. Moreover, we use here an algebraic definition of pointedness, although it does have a geometrical interpretation in the all-contact case: the friction cone $D_{\mathscr{S}}\Upsilon$ does not contain a nontrivial linear space, and thus its origin is "pointed". The reason we use the algebraic formulation is that it applies for cases where joint constraints are also included and where the constraint set is algebraically pointed, but not geometrically. More importantly the algebraic definition is also intuitive from a mechanics perspective: there are no contact forces that are valid from a constraint perspective but that produce a zero total force. In other words, the multibody system cannot get stuck through its internal forces, a configuration known to lead to virtually unpredictable behavior. For one contact, this condition immediately holds when we have rolling, sliding, and spinning friction; for more contacts, whether this holds depends on the configuration.

With this definition we have the following result, which elucidates the existence and uniqueness of a solution.

Theorem 2 *Assume that the friction cone attached to the CCP (or VI) (41) is pointed in the sense of definition (45). Then*

(i) *The VI (41) has a solution.*
(ii) *Any two solutions of the VI result in the same velocity vector (44).*

Proof If the friction cone is pointed, then, by Theorem 1(ii), the optimization problem (42) must have a solution. Then, by Theorem 1, that solution is also a solution of the variational inequality (41), which proves the claim (i). For the second part, assume that the VI (41) has two solutions, $\boldsymbol{\gamma}^{1}_{\mathscr{S}}$, and $\boldsymbol{\gamma}^{2}_{\mathscr{S}}$. From Theorem 1(i), these are also solutions of the optimization problem (42). From Theorem 1(ii), these solutions satisfy

$$D_{\mathscr{S}}\boldsymbol{\gamma}^{1}_{\mathscr{S}} = D_{\mathscr{S}}\boldsymbol{\gamma}^{2}_{\mathscr{S}}.$$

Looking at (44) and how the velocity solutions are computed from the solutions of (42), one can see that this relationship implies that the velocities computed from either are the same. The proof of (ii) and the theorem is complete. □

With this theorem, it follows that under the condition of pointedness of the friction cone, the VI (41) is well posed and its solution is unique in velocities. The latter condition is sufficient to result in a predictive simulation, since it results in uniqueness of the trajectory. We also point out that in the case of friction it is unreasonable to expect that the forces are unique, as can be immediately contemplated from the example of a block resting on a table with friction, which has multiple frictional solutions.

5 Numerical solution scheme

Obtaining the unknown $\boldsymbol{\gamma}_{\mathscr{S}}$ dual variables from the VI of Eq. (41) is the most complex and time-consuming task in the simulation process. In fact, the solution of large VIs is currently a debated and actively researched topic in applied mathematics, and there exist no unique "best" algorithms to approach their solution.

As a comparison, the much easier linear problems that often arise in the simulation of classical ODEs and DAEs belong to the so-called polynomial **P**-class because they can be solved in polynomial time $O(n^c)$, if n represents the size of the problem in our case, the number of constraints on the system. On the other hand, VIs (as well as QPs, LCPs, and CCPs as special cases) are highly nonlinear problems, whose general complexity class is said to be nondeterministic-polynomial **NP**-hard, which often means an intractable growth of computational time for even moderate sizes of the domain [4]. Only special subcases of VIs can exploit the more friendly and tractable **P** class, such as the monotone VIs that we are dealing with. Exact methods for their solution, although running with polynomial time $O(n^c)$, might experience high c; thus the solution is still difficult, and often it is necessary to accept truncation to approximate solutions [22].

In [33] we presented a fixed-point iteration that can solve the CCP problem with Amontons-Coulomb friction only. Adding the set-valued rolling and spinning friction model presented in this paper, one can develop a new flavor of that iterative scheme, with minor modifications. Omitting the details, we note that the method iterates the following function, which is convergent for a proper choice of $\omega \in \mathbb{R}^+$, $\lambda \in (0, 1]$ and K:

$$\begin{aligned}\boldsymbol{\gamma}_{\mathscr{S}}^{r+1} = \lambda \Pi_{\Upsilon}\big(\boldsymbol{\gamma}_{\mathscr{S}}^{r} - \omega B^r \big(N\boldsymbol{\gamma}_{\mathscr{S}}^{r} + \mathbf{r} \\ + K^r\big(\boldsymbol{\gamma}_{\mathscr{S}}^{r+1} - \boldsymbol{\gamma}_{\mathscr{S}}^{r}\big)\big)\big) + (1-\lambda)\boldsymbol{\gamma}_{\mathscr{S}}^{r}, \qquad (46)\\ r = 0, 1, 2, \ldots.\end{aligned}$$

The Π_{Υ} operator is a projection onto Υ, so that $\Pi_{\Upsilon}(\boldsymbol{\gamma}) = \operatorname{argmin}_{\zeta\in\Upsilon}\|\boldsymbol{\gamma} - \zeta\|$. Given the separable cone structure of Π_{Υ} as defined in Eq. (35), the computation of $\Pi_{\Upsilon}(\boldsymbol{\gamma})$ can be split into $n_{\mathscr{A}}$ projections $\Pi^i_{\mathscr{Z}}(\boldsymbol{\gamma}^i)$ and into $n_{\mathscr{B}}$ identities (no projections are required for the bilateral constraints). The original $\Pi^i_{\mathscr{Z}}(\boldsymbol{\gamma}^i)$ projection discussed in [33] operates only on the three values of the reactions that must obey the Amontons-Coulomb friction. In the advanced case, however, that includes also rolling friction and spinning friction one must map the six-dimensional wrench onto the six-dimensional cone $\mathscr{Z}$ defined in Eq. (31), so $\Pi^i_{\mathscr{Z}}(\boldsymbol{\gamma}^i) : \mathbb{R}^6 \to \mathbb{R}^6$. See Appendix C for details about this projection.

Given the requirement of solving a VI for each timestep, the DVI approach appears to be less competitive than the classic DEM concept of regularizing nonsmooth phenomena via smooth but stiff force fields, because DEMs lead to uncoupled equations of motion of ODE type with basic linear $O(n_b)$ complexity for each timestep. However, our DVI setting permits large timesteps because it does not suffer from the stability issues implied by stiff forces in DEMs, so the increased workload for each timestep is paid back by fewer timesteps being required [32].

We remark that a parallel version of this rolling and spinning friction model can be easily implemented on parallel hardware of GPU type, as described in [34], and on hybrid high-performance computers [21].

6 Examples

6.1 Comparison against analytical solution

A simple validation against the analytical solution for a rolling disk is presented here. We consider a rolling disk with radius $R = 1$ m, mass $M = 10$ kg, moment of inertia $J = 4\,\mathrm{kg\,m^2}$, initial position $x|_{t=0} = 0$ m initial horizontal speed on x axis $v_x|_{t=0} = 1$ m/s, initial angular velocity $\omega|_{t=0} = -1$ rad/s. We use a rolling friction parameter $\rho = 0.02$ m and a sliding/static friction coefficient $\mu = 0.9$. The gravity acceleration is $g_y = -9.8\,\mathrm{m/s^2}$.

Figure 2 plots the motion of the disk which rolls until it comes to a rest because of the deceleration caused by the rolling friction. Figure 3 shows that the deceleration is constant (linear speed) in analytical theory, which uses the constant resistant torque $T_r = M \cdot g_y \cdot \rho$; a similar behavior can be seen in the results of the DVI model.

6.2 Simulation of a linear guideway

Linear guideways based on recirculating ball bearings represent an advanced technology for constraining linear motion. They avoid the problem of sliding friction between the translating parts by interposing

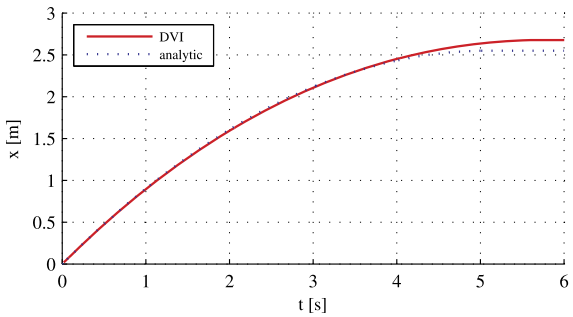

Fig. 2 Motion of the rolling disk: comparison with classic theory

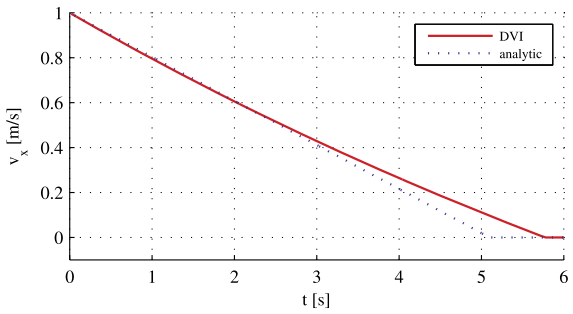

Fig. 3 Velocity of the rolling disk: comparison with classic theory

rows of rolling spheres that can recirculate back in enclosed channels, like tracks in a tank. In this section we present the simulation of the linear system depicted in Fig. 4, as an example of a device that involves many contacts with rolling friction. The model is based on a 35 mm wide rail, with four rows of recirculating balls, each steel ball having a diameter of $2R_b = 6.4$ mm. Nine balls, on average, provide the contact in each row, and the four rows define four contact directions with 45° orientation, as shown in Fig. 5.

Although this type of guideway avoids sliding friction, a small resistance remains, caused by rolling friction, as shown in the scheme of Fig. 6. This effect can be easily simulated by using the proposed complementarity-based approach to rolling friction: we modeled the guideway in a 3D CAD software, we saved the parts using a custom translation software, we imported the file in our software, and we simulated it.

We assumed that the contact between the spheres and the raceways is always aligned to the 45° direction, so we introduced eight stretched boxes as collision shapes, as shown in Fig. 7; in this way the collision algorithm of our software can automatically com-

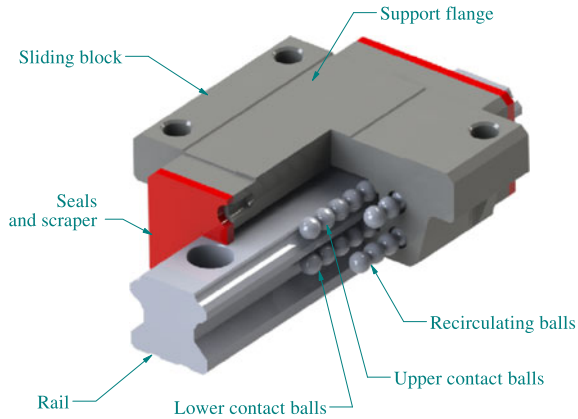

Fig. 4 The simulated linear guideway, with the scheme of the recirculating balls

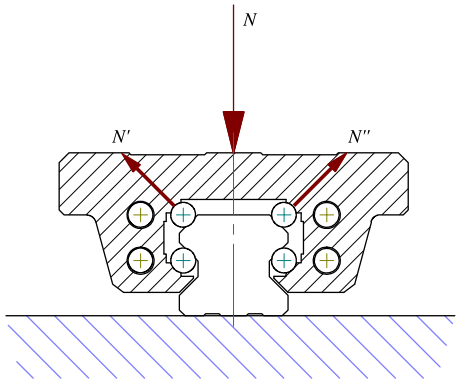

Fig. 5 Section of the linear guideway, showing the contact between balls and grooves

pute the contact points between the balls and the raceways. In other types of guideways, the raceways are shaped like Gothic arcs: these can be modeled as well, but such nodeling would not add much to the discussion.

No significant side contact occurs between rolling balls (in some models plastic spacers exclude this possibility), and we did not consider static preload, although it could be simulated as well.

Since the system has many more contact constraints than needed, the indeterminacy is solved by introducing a Tikhonov regularization: from a numerical point of view this means adding a nonzero diagonal C to the N matrix of Eq. (38), that is, $N_{\mathrm{Tyk}} = N + C$. From a mechanical point of view, the Tikhonov regularization means that we introduce compliance in contacts; in detail, we generate C from the inverses of the stiffness

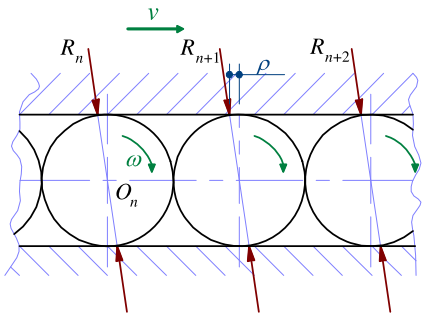

Fig. 6 Schematic representation of the forces acting on the rolling balls

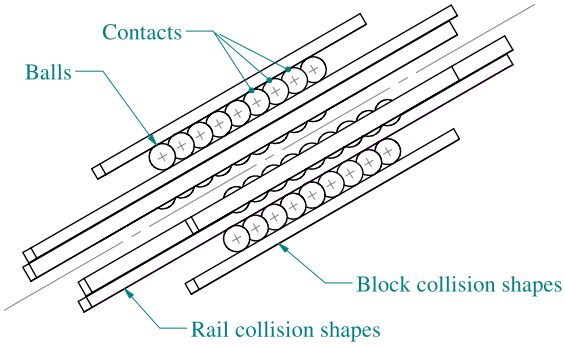

Fig. 7 Collision primitives used in the simulation

values in contact points, as computed by the Hertz-Mindlin theory.

The manufacturer of the guideway provides the following formula for estimating the resistance to horizontal sliding at low speeds: $F_x = -\text{sign}(v_x)(\zeta F_y + S)$, where F_y is the normal load, ζ is typically about 0.005, and S is the force caused by the sliding friction of seals and scrapers, in our case $S = 5.3$ N. Such friction is introduced in our model by using a convex box-constraint of the type $-S \le \hat{\gamma}_S \le S$, whose effect on the DVI is similar to the already discussed, and more complex, contact constraints. The effect of the ζF_y term comes from the simulation of the many rolling contacts, each with a rolling friction parameter $\rho = \zeta R_b$. Results from the simulations show precise agreement with the above formula, as shown in Fig. 8. For zero speed, the model is able to describe also the sticking effect that, although modest, can be measured on this class of devices.

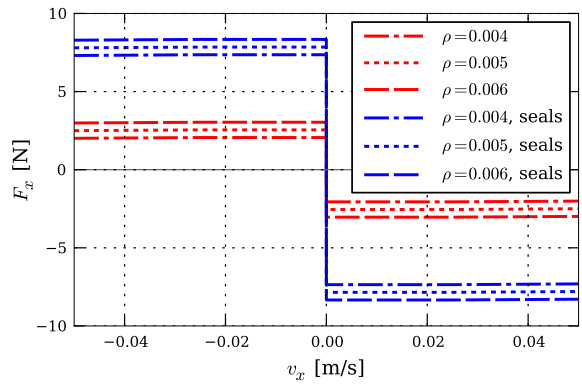

Fig. 8 Resisting force F_x for different values of rolling friction parameter ρ, with and without the effect of seal friction (case of load $F_y = 490$ N)

6.3 Application to the simulation of granular materials

The mechanics of granular matter has became a fertile research topic only recently, because of the vast computational resources that are required. One of the fields that would benefit from advances in this area is pharmaceutical engineering, where multibody dynamics could be used to study processes that involve powders: milling, blending, granulation, compression, and coating [17].

In the popular discrete element method (DEM), the bulk material is discretized in many particles with unilateral frictional contacts [7]. Various micromechanical contact models are available in the DEM field in order to define the interaction between the particles; in most cases, the nonsmooth nature of contact means that those models always produce stiff contact forces, and hence, that short integration timesteps are needed.

The need of rolling friction in granular simulations is motivated by experimental evidence; for example [23] shows that rolling resistance can have marked influence on the mechanics of particle assemblies at microscales; in some cases, its effect can be more relevant than interparticle sliding friction [36]. Rolling friction has been shown to affect only marginally the elastic properties of granular assemblies, but other collective phenomena such as shear resistance and dilatancy are significantly affected [5].

In some cases the simple inhibition of particle rotations in DEM algorithms can improve the solution with respect to the case of free frictionless rotation [6]. During the past few years, more sophisticated models

Fig. 9 A typical pile of gravel and a conveyor, in a plant for separating raw materials

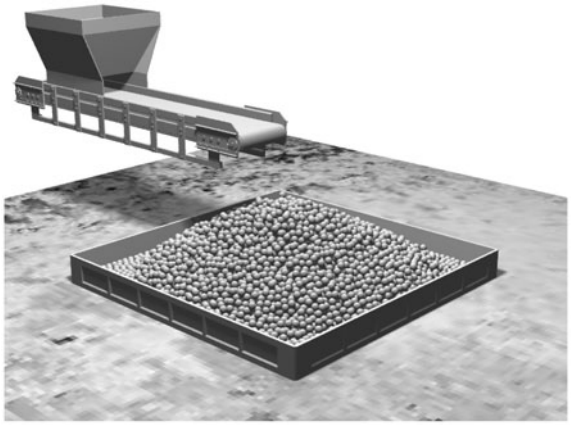

Fig. 10 Simulated pile of gravel, without rolling friction. Case with $\rho = 0$ m, $\sigma = 0$ m, after 13 s of simulation

of rolling friction have been proposed for interparticle contacts, for example in [14]. Given the difficulty of tuning the parameters of complex models, approaches based on few parameters such as the approach of [15] are welcome.

To some extent, the macroscopic effect of rolling friction in granular media is similar to the effect of dealing with nonspherical particles [19].

The collective behavior of particles with irregular and faceted shapes is different from the behavior of spherical particles, even if granular assemblies share the same granulometry and friction; in general, oddly shaped particles tend to generate less deformable assemblies when compared with spherical particles of equal size [35]. Of course, a straightforward approach could take into account the simulation of all the detailed shapes, but this would lead to high simulation times, both because there will be multiple contacts between pairs of particles and because the collision detection phase would require more RAM and CPU time to process those contacts.

The rolling friction model discussed above can be used for simulating the granular materials such as in the example of Fig. 9. If one tries to simulate the pile of gravel with plain rigid spheres, the angle of repose of the cone will be small when compared with the real case, because the lack of irregularities on the simulated spheres lead to a loose, granular flow. Yet, by introducing increasing values of rolling and spinning friction in a model with simple spheres, the stacking is less loose and we can obtain the same results that one can achieve by introducing particles with odd shapes (which would require much larger computational resources).

We simulated the free stacking of 10,000 particles, each with a diameter $D_s = 45$ mm, falling from an height of 1 m into a 2 m $\times$ 2 m flat container. The density of the particles is $\delta = 2\,028$ kg/m^2; the sliding friction coefficient is $\mu = 0.6$, also used as static friction coefficient. The flow is about 800 particles/s, and gravity is $g = -9.81$ m/s^2. We simulated this system for increasing values of rolling and spinning friction. Specifically, we tested values of rolling friction parameter in the 0–0.01 m range; for simplicity we made the spinning friction parameter equal to the rolling one in all tests. The timestep was $h = 0.005$ s, and the simulated time 15 s. As shown in Fig. 10, the case without rolling friction produces an almost flat stack, whereas Fig. 11 shows that the proposed rolling friction model is able to produce a steep cone typical of particles of irregular shapes yet featuring the benefit of using simple spheres (see also Figs. 12 and 13).

7 Conclusion

We presented a rolling friction model that fits in the context of DVIs, and we discussed theoretical issues about the existence and uniqueness of the solution. Details about an implementation in the form of a convex cone-complementarity problem are given, showing that the approach features a relatively simple algorithmic complexity yet provides stable, fast, and robust solutions. These features, coupled with the ease of use, make the method a good candidate for enriching the

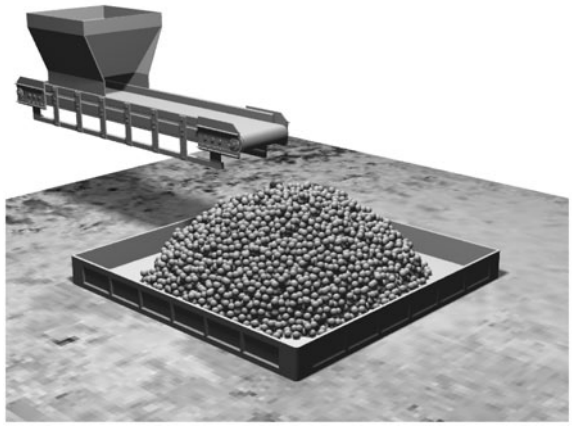

Fig. 11 Simulated pile of gravel, with rolling friction $\rho = 0.01$ m and spinning friction $\sigma = 0.01$ m

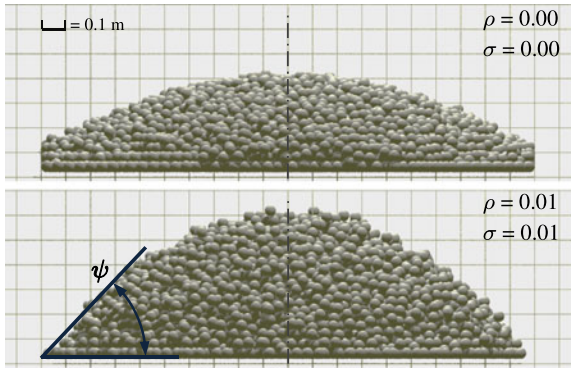

Fig. 12 Angle of repose for different values of rolling friction

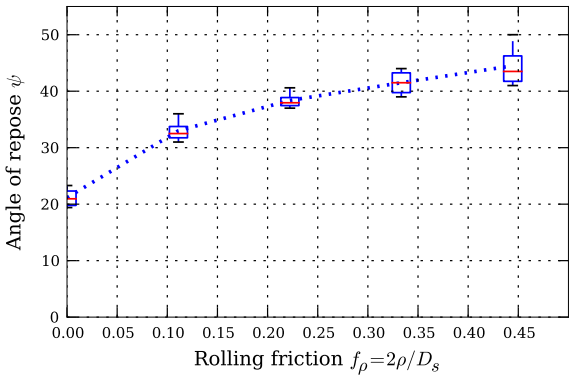

Fig. 13 Angle of repose ψ with same $\mu = 0.6$, for increasing values of ρ

capabilities of DVI simulations of rigid contacts, especially when fast, real-time performance is required or when a large number of parts is involved, such as in granular flows.

Acknowledgements A. Tasora thanks Ferrari Automotive and TP Engineering for financial support. Mihai Anitescu was supported by the U.S. Department of Energy, under Contract No. DE-AC02-06CH11357.

Appendix A: Kinematics of rolling in three-dimensional space

Let $\omega^{(W)}_{A,W}$ and $\omega^{(W)}_{B,W}$ denote the angular velocities of two bodies A and B, relative to the absolute reference frame W and expressed in the basis of the frame (W). We assume A and B to be rigid or with negligible deformations. Introducing the rotation matrix $R_{A,W} \in \mathsf{SO}(\mathbb{R}, 3)$ that represents the rotation of A respect to W, we have $\omega^{(W)}_{A,W} = R_{A,W}\omega^{(A)}_{A,W}$ and $\omega^{(W)}_{B,W} = R_{B,W}\omega^{(A)}_{B,W}$.

Let the unimodular quaternion $\varepsilon_{A,W} \in \mathbb{H}_1$ represent the rotation of the frame A respect to absolute frame W. We recall that, for unimodular quaternions, the inverse ε^{-1} is also the conjugate ε^*. We also recall that it is possible to compute R from ε and vice versa.

Thanks to a property of quaternion algebra [13] the relative rotation of two references is

$$\varepsilon_{B,A} = \varepsilon^*_{A,W}\varepsilon_{B,W}.$$

By performing differentiation respect to time, we get

$$\dot{\varepsilon}_{B,A} = \dot{\varepsilon}^*_{A,W}\varepsilon_{B,W} + \varepsilon^*_{A,W}\dot{\varepsilon}_{B,W}.$$

From the result in [27], the quaternion derivative can be transformed in angular velocity, using *pure* quaternions:

$$\begin{aligned}\left[0, \omega^{(A)}_{BA}\right] &= 2\dot{\varepsilon}_{B,A}\varepsilon^*_{B,A} \qquad (47)\\ &= 2\dot{\varepsilon}^*_{A,W}\varepsilon_{B,W}\left(\varepsilon^*_{A,W}\varepsilon_{B,W}\right)^* \\ &\quad + 2\varepsilon^*_{A,W}\dot{\varepsilon}_{B,W}\left(\varepsilon^*_{A,W}\varepsilon_{B,W}\right)^*. \qquad (48)\end{aligned}$$

Since $(\varepsilon^*_1\varepsilon_2)^* = \varepsilon^*_2\varepsilon_1$, and remembering that $\varepsilon\varepsilon^* = 1$, we can develop Eq. (48) into

$$\left[0, \omega^{(A)}_{BA}\right] = 2\dot{\varepsilon}^*_{A,W}\varepsilon_{A,W} + 2\varepsilon^*_{A,W}\dot{\varepsilon}_{B,W}\varepsilon^*_{B,W}\varepsilon_{A,W}. \qquad (49)$$

The product $\dot{\varepsilon}_{B,W}\varepsilon^*_{B,W}$ in the second term of the summation can be replaced with the pure quaternion $\frac{1}{2}[0, \omega^{(W)}_{B,W}]$ using Eq. (47). Also, the first term

can be premultiplied by $\varepsilon^*_{A,W}\varepsilon_{A,W} = 1$, becoming $2\varepsilon^*_{A,W}\varepsilon_{A,W}\dot{\varepsilon}^*_{A,W}\varepsilon_{A,W}$; here the product between the second and third quaternion can be replaced with the pure quaternion $\frac{1}{2}[0, \omega^{(W)}_{A,W}]^*$, again using Eq. (47). Thus we have

$$\begin{aligned}\big[0, \omega^{(A)}_{BA}\big] &= \varepsilon^*_{A,W}\big[0, \omega^{(W)}_{A,W}\big]^*\varepsilon_{A,W} \\ &\quad + \varepsilon^*_{A,W}\big[0, \omega^{(W)}_{B,W}\big]\varepsilon_{A,W}. \end{aligned} \tag{50}$$

A rotation in 3D space of the vector part of a pure quaternion can be obtained with unitary quaternions, that is, $[0, \mathbf{v}^{(W)}] = \varepsilon_{A,W}[0, \mathbf{v}^{(A)}]\varepsilon^*_{A,W}$.

Hence, recalling that $[0, \omega^{(W)}_{A,W}]^* = -[0, \omega^{(W)}_{A,W}]$ by the property of conjugate quaternions, we can rewrite Eq. (50) and obtain the expected result for relative angular velocity ω_r:

$$\begin{aligned}\big[0, \omega^{(A)}_{BA}\big] &= -\big[0, \omega^{(A)}_{A,W}\big] + \big[0, \omega^{(A)}_{B,W}\big] \\ \omega^{(A)}_{BA} &= \omega^{(A)}_{B,W} - \omega^{(A)}_{A,W}. \end{aligned} \tag{51}$$

Appendix B: Formulation of D vectors

We assume that the vector of generalized velocities $\mathbf{v}$ contains the speeds of the centers of mass of the bodies $\dot{\mathbf{x}}^{(W)}$, expressed in absolute coordinates (W) and the angular velocities $\omega^{(i)}$ expressed in the local coordinates of the ith body, as $\mathbf{v} = [\dot{\mathbf{x}}^{(W)}_1, \omega^{(1)}_1, \dot{\mathbf{x}}^{(W)}_2, \omega^{(2)}_2, \ldots]^T$.

Given a contact between a pair of two rigid bodies A and B, we define the positions of the two contact points with respect to the centers of mass, expressed in the coordinate systems of the two bodies, as $\mathbf{s}^{(A)}_A$ and $\mathbf{s}^{(B)}_B$. The absolute rotations of the coordinate systems of the bodies are $R^{(W)}_A, R^{(W)}_B \in \mathsf{SO}(\mathbb{R}, 3)$ and the absolute rotation of the contact plane is $R^{(W)}_P \in \mathsf{SO}(\mathbb{R}, 3) = [\mathbf{n}, \mathbf{u}, \mathbf{w}]$. Thus, the vectors $\mathbf{D}_{\gamma_n}, \mathbf{D}_{\gamma_u}, \mathbf{D}_{\gamma_w}$ can be computed as $D_\gamma = [\mathbf{D}_{\gamma_n}, \mathbf{D}_{\gamma_u}, \mathbf{D}_{\gamma_w}] \in \mathbb{R}^{3\times m_v}$,

$$\begin{aligned} D^T_\gamma = \big[0, &\quad \ldots, \quad R^{(W)^T}_P, \quad -R^{(W)^T}_P R^{(W)}_A \tilde{s}^{(A)}_A, \quad \ldots, \\ &0, \quad \ldots, \quad -R^T_P, \quad R^{(W)^T}_P R^{(W)}_B \tilde{s}^{(B)}_B, \\ &\ldots, \quad 0\big], \end{aligned} \tag{52}$$

where $\tilde{s}$ is the skew symmetric matrix such that $\tilde{s}\mathbf{x} = \mathbf{s} \wedge \mathbf{x}$.

Similarly, recalling the result in Eq. (51), one can compute the vectors $\mathbf{D}_{\tau_n}, \mathbf{D}_{\tau_u}, \mathbf{D}_{\tau_w}$ as $D_\tau = [\mathbf{D}_{\tau_n}, \mathbf{D}_{\tau_u}, \mathbf{D}_{\tau_w}] \in \mathbb{R}^{3\times m_v}$:

$$\begin{aligned} D^T_\tau = \big[0, &\quad \ldots, \quad 0, \quad R^{(W)^T}_P R^{(W)}_A, \quad \ldots, \\ &0, \quad \ldots, \quad 0, \quad -R^{(W)^T}_P R^{(W)}_B, \quad \ldots, \quad 0\big]. \end{aligned} \tag{53}$$

We remark that, because of the extreme sparsity of (52) and (53), only the following four 3×6 matrices need to be stored per each contact

$$D^T_{\gamma,A} = \big[R^{(W)^T}_P, \quad -R^{(W)^T}_P R^{(W)}_A \tilde{s}^{(A)}_A\big] \tag{54}$$

$$D^T_{\gamma,B} = \big[-R^{(W)^T}_P, \quad R^{(W)^T}_P R^{(W)}_B \tilde{s}^{(B)}_B\big] \tag{55}$$

$$D^T_{\tau,A} = \big[0, \quad R^{(W)^T}_P R^{(W)}_A\big] \tag{56}$$

$$D^T_{\tau,B} = \big[0, \quad -R^{(W)^T}_P R^{(W)}_B\big] \tag{57}$$

Here we considered B as the reference body: otherwise, if A were the reference for contact coordinates, signs should be swapped in all terms in Eqs. (52)–(57).

Appendix C: Computing projections on intersections of cones

We describe the procedure to compute the Euclidean projection of a point x on an intersection of circular cones that have one common component (in the case studied here, that component is the normal force). We assume that a generic point x is structured as follows:

$$x = (x_0, l_1, l_2, \ldots, l_m), \quad x_0 \in \mathbb{R},\ l_i \in \mathbb{R}^{n_i}, \tag{58}$$

and that the m circular cones are second-order cones defined by

$$x \in K_i \quad \Leftrightarrow \quad \mu_i x_0 \geq \sqrt{\|l_i\|^2},$$

where $\mu_i > 0$, $i = 1, 2, \ldots, m$. We are interested in computing the projection of a vector x on $\bigcap K_i$, that is,

$$\tilde{x} = \prod_{\bigcap K_i}(x) \quad \Leftrightarrow \quad \|x - \tilde{x}\|^2 = \min_{y \in \bigcap K_i} \|x - y\|^2.$$

For example, in the case treated in this work, we are interested in simultaneous modeling of sliding, rolling, and spinning friction in three dimensional configurations. That is, we have three cones, $m = 3$ and x is a six-dimensional vector, $x = (\gamma_n, \gamma_u, \gamma_w, \tau_u, \tau_w, \tau_n)$. The mapping (58) is the following: $x_0 = \gamma_n$, $l_1 = (\gamma_u, \gamma_w)$, $l_2 = (\tau_u, \tau_w)$, $l_3 = \tau_n$. The friction coefficients are $\mu_1 = \mu$, $\mu_2 = \rho$, $\mu_3 = \sigma$.

The crucial observation that simplifies the computation of the projection is that the component $\tilde{l}_i$ of the

projection $\tilde{x}$ must be collinear with l_i. Indeed, if this is not the case, then rotating $\tilde{l}_i$ over l_i will preserve feasibility but will necessarily reduce $\|x - \tilde{x}\|$, a contradiction. Therefore, there exists t_i such that $\tilde{l}_i = t_i l_i$. The optimization that defines the projection then becomes

$$\min_{y_0,t_1,t_2,\ldots,t_m} (y_0 - x_0)^2 + \sum_{i=1}^{m} \left(t_i \frac{\mu_i y_0}{\|l_i\|} - 1 \right)^2 \|l_i\|^2,$$
$$0 \le t_i \le 1,\ i = 1, 2, \ldots, m.$$

We have normalized the component of y in terms of y_0 to allow for the range of t_i to be the same. For a given y_0, the optimal t_i, which we denote by $t_i(y_0)$, is easy to compute. Indeed we obtain the following

$$t_i(y_0) = \begin{cases} \frac{\|l_i\|}{\mu_i y_0} & \frac{\|l_i\|}{\mu_i y_0} \le 1 \\ 1 & \frac{\|l_i\|}{\mu_i y_0} > 1 \end{cases}$$

$$\Rightarrow \quad t_i(y_0) \frac{\mu_i y_0}{\|l_i\|} - 1 = \begin{cases} 0 & \frac{\|l_i\|}{\mu_i y_0} \le 1 \\ \frac{\mu_i y_0}{\|l_i\|} - 1 & \frac{\|l_i\|}{\mu_i y_0} > 1 \end{cases}$$

Substituting t_i for the optimal values $t_i(y_0)$ in the optimization problem, we obtain that the problem is equivalent to

$$\min_{y_0} \psi(y_0) := (y_0 - x_0)^2 + \sum_{i=1}^{m} I_{[y_0 < \frac{\|l_i\|}{\mu_i}]}(y_0) \left(\frac{\mu_i y_0}{\|l_i\|} - 1 \right)^2 \|l_i\|^2.$$

Here I is the indicator function of a set. It is immediately apparent that this function is piecewise quadratic and *that it is convex*. Indeed, convexity follows from the fact that each term function is convex, the first term as a quadratic, and the other terms as their graphs are the union of a parabola with a flat line.

To find its optimum, we can do the following.

1. Define and order the breakpoints 0, and $\frac{\|l_i\|}{\mu_i}$, with $i = 1, 2, \ldots, m$. Successive breakpoints define a piece.
2. On each piece find the minimum of the quadratic function.
3. Compute the overall minimum, which is the lowest value of all such minima.

Once $\tilde{x}_0 = y_0$ is determined, $t_i(y_0)$ is computed, and the other components of the projection are computed as $\tilde{l}_i = t_i(\tilde{x}_0) \frac{\mu_i \tilde{x}_0}{\|l_i\|}$.

For a large number of breakpoints we can exploit convexity of ψ, by noting that we can evaluate the function at the breakpoints, and find the minimum value. Then, by convexity, the overall minimum must occur in a segment that neighbors the breakpoint with the minimum value. Hence, one minimizes the quadratic only in those intervals.

To summarize:

1. Define and order the breakpoints 0, and $\frac{\|l_i\|}{\mu_i}$, with $i = 1, 2, \ldots, m$. Successive breakpoints define a piece. We assume without loss of generality that the labels have been permuted so that the natural order has the breakpoints in increasing order, that is, $i < j \Rightarrow \frac{\|l_i\|}{\mu_i} < \frac{\|l_j\|}{\mu_j}$. If two breakpoints have the same value, we delete their index.
2. Enumerate the objective function ψ at the breakpoints, and find the i for which $\psi(\frac{\|l_i\|}{\mu_i}) \le \psi(\frac{\|l_j\|}{\mu_j})$, $\forall j$. If there is one such i, the overall minimum is on a neighboring segment; if there are two, it is on the segment in between (there cannot be three different indices, since the function is not piecewise constant).
3. Minimize the piecewise quadratic on either the one or two segments identified, and report the result.

For a small number of breakpoints (i.e., the number of cones m is small), it is not likely that this reduced method would practically be much faster than comprehensive enumeration.

References

1. Anitescu M, Hart GD (2004) A fixed-point iteration approach for multibody dynamics with contact and friction. Math Program, Ser B 101(1):3–32 (ANL/MCS P985-0802)
2. Anitescu M, Tasora A (2010) An iterative approach for cone complementarity problems for nonsmooth dynamics. Comput Optim Appl 47(2):207–235
3. Anitescu M, Potra FA, Stewart D (1999) Time-stepping for three-dimensional rigid-body dynamics. Comput Methods Appl Mech Eng 177:183–197
4. Arora S, Barak B (2009) Computational complexity: a modern approach. Cambridge University Press, Cambridge
5. Bardet J (1994) Observations on the effects of particle rotations on the failure of idealized granular materials. Mech Mater 18(2):159–182. Special Issue on Microstructure and Strain Localization in Geomaterials
6. Calvetti F, Nova R (2004) Micromechanical approach to slope stability analysis. Degradations and instabilities in geomaterials. Springer, Berlin

7. Cundall PA, Strack ODL (1979) A discrete numerical model for granular assemblies. Geotechnique 29(1):47–65
8. de Coulomb CA (1821) Théorie des machines simples en ayant égard au frottement de leurs parties et à la roideur des cordages. Bachelier, Paris
9. Estrada N, Azema É, Radjaï F, Taboada A (2011) Identification of rolling resistance as a shape parameter in sheared granular media. 1. Phys Rev E, Stat Nonlinear Soft Matter Phys 84(1):011306
10. Facchinei F, Pang J (2003) Finite-dimensional variational inequalities and complementarity problems, vol 1. Springer, Berlin
11. Flores P, Leine R, Glocker C (2012) Application of the nonsmooth dynamics approach to model and analysis of the contact-impact events in cam-follower systems. Nonlinear Dyn 69:2117–2133. doi:10.1007/s11071-012-0413-3
12. Hairer E, Nørsett SP, Wanner G (2010) Solving ordinary differential equations. Springer, Berlin
13. Haug EJ (1989) Computer-aided kinematics and dynamics of mechanical systems. Prentice-Hall, Englewood Cliffs
14. Iwashita K, Oda M (1998) Rolling resistance at contacts in simulation of shear band development by DEM. J Eng Mech 124(3):285–292
15. Jiang M, Yu H-S, Harris D (2005) A novel discrete model for granular material incorporating rolling resistance. Comput Geotech 32(5):340–357
16. Jourdan F, Alart P, Jean M (1998) A Gauss Seidel like algorithm to solve frictional contract problems. Comput Methods Appl Mech Eng 155:31–47
17. Ketterhagen WR, am Ende MT, Hancock BC (2009) Process modeling in the pharmaceutical industry using the discrete element method. J Pharm Sci 2(98):442–470
18. Kinderleher D, Stampacchia G (1980) An introduction to variational inequalities and their application. Academic Press, New York
19. Kruggel-Emden H, Rickelt S, Wirtz S, Scherer V (2008) A study on the validity of the multi-sphere discrete element method. Powder Technol 188(2):153–165
20. Leine RI, Glocker C (2003) A set-valued force law for spatial Coulomb-Contensou friction. Eur J Mech A, Solids 22(2):193–216
21. Negrut D, Tasora A, Mazhar H, Heyn T, Hahn P (2012) Leveraging parallel computing in multibody dynamics. Multibody Syst Dyn 27:95–117. doi:10.1007/s11044-011-9262-y
22. Nocedal J, Wright SJ (1999) Numerical optimization, vol. 39. Springer, Berlin
23. Oda M, Konishi J, Nemat-Nasser S (1982) Experimental micromechanical evaluation of strength of granular materials: effects of particle rolling. Mech Mater 1(4):269–283
24. Pacejka HB (2005) Tire and vehicle dynamics, 2nd edn. SAE International, Warrendale
25. Pfeiffer F, Glocker C (1996) Multibody dynamics with unilateral contacts. Wiley, New York
26. Rankine WJM (1868) Manual of applied mechanics. Charless Griffin, London
27. Shabana AA (2005) Dynamics of multibody systems, 3rd edn. Cambridge University Press, Cambridge
28. Stewart D, Pang J-S (2008) Differential variational inequalities. Math Program 113(2):345–424
29. Stewart DE (2001) Reformulations of measure differential inclusions and their closed graph property. J Differ Equ 175:108–129
30. Stewart DE, Trinkle JC (1996) An implicit time-stepping scheme for rigid-body dynamics with inelastic collisions and Coulomb friction. Int J Numer Methods Eng 39:2673–2691
31. Studer C, Glocker C (2007) Solving normal cone inclusion problems in contact mechanics by iterative methods. J Syst Des Dyn 1(3):458–467
32. Tasora A, Anitescu M (2010) A convex complementarity approach for simulating large granular flows. J Comput Nonlinear Dyn 5(3):1–10
33. Tasora A, Anitescu M (2011) A matrix-free cone complementarity approach for solving large-scale, nonsmooth, rigid body dynamics. Comput Methods Appl Mech Eng 200(5–8):439–453
34. Tasora A, Negrut D, Anitescu M (2008) Large-scale parallel multi-body dynamics with frictional contact on the graphical processing unit. J Multi-Body Dyn 222(4):315–326
35. Terzaghi K, Peck RB, Mesri G (1996) Soil mechanics in engineering practice. Wiley-Interscience, New York
36. Tordesillas A, Walsh D (2002) Incorporating rolling resistance and contact anisotropy in micromechanical models of granular media. Powder Technol 124(1–2):106–111
37. Weisbach JL (1870) A manual of the mechanics of engineering and of the construction of machines, vol 3. Van Nostrand, New York