# On enhanced descent algorithms for solving frictional multicontact problems: application to the discrete element method 

Serge Dumont

## - To cite this version:

Serge Dumont. On enhanced descent algorithms for solving frictional multicontact problems: application to the discrete element method. International Journal for Numerical Methods in Engineering, 2013, 93 (11), pp.1170-1190. 10.1002/nme. 4424 . hal-01113707

# HAL Id: hal-01113707 

## https://hal.science/hal-01113707

Submitted on 10 Mar 2023

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. publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

# On enhanced descent algorithms for solving frictional multicontact problems: application to the discrete element method 

S. Dumont ${ }^{*} \dagger$<br>LAMFA, Université de Picardie Jules Verne - CNRS UMR 7352, 33, rue Saint-Leu, France


#### Abstract

In this article, various numerical methods to solve multicontact problems within the nonsmooth discrete element method are presented. The techniques considered to solve the frictional unilateral conditions are based both on the bipotential theory introduced by G. de Saxcé and the augmented Lagrangian theory introduced by P. Alart. Following the ideas of Z.-Q. Feng a new Newton method is developed to improve these classical algorithms, and numerical experiments are presented to show that these methods are faster than the previous ones, provide results with a better quality, and are less sensitive to the numerical parameters. Moreover, a stopping criterion that ensures a good mechanical property of the solution is provided.


KEY WORDS: granular materials; contact mechanics; Newton algorithms; bipotential; augmented Lagrangian

## 1. INTRODUCTION

The simulation of granular media concerns a wide range of engineering applications, such as concrete or geomaterials. All these materials can be considered as particles in interaction. In this study, the interaction is the frictional contact.

The numerical simulation of frictional contact problem remains a difficult problem, because the frictional contact law is strongly nonlinear and multivalued. Nevertheless, several approaches exist to solve contact problems. A wide bibliography exists on this subject, and an extensive list of references can be found, for example, in [4] or in [5]. The most popular method is based on the penalty approximation, due to its simplicity. But its important drawback is that the numerical parameters that have to be considered are difficult to choose efficiently. Indeed, if the penalty coefficient is small, then the contact laws are not accurately satisfied, and if the penalty coefficient is larger, then the problem to be solved becomes stiff.

Another important family of approaches is based on the notion of augmented Lagrangian. But because the frictional contact problem is no longer a classical minimization problem, the standard approaches need to be extended. This article deals with two methods of this family, which are the Alart and Curnier [2] method and the De Saxcé and Feng [1] method. These methods have been already adapted to the solving of multicontact problems, in the LMGC90 software for the first method and in the MULTICOR software for the second.

More precisely, the Alart and Curnier method consists in a Newton method to find the saddle point of an augmented Lagrangian method where the convex of constraints depends on the solution.

[^0]The method developed by De Saxcé and Feng is based on the theory of implicit standard materials and the writing of a so-called bipotential, which is minimized using a gradient method.

In this article, we will consider the works of Feng et al. [3] to improve the resolution of the multibody problem, firstly, within the bipotential framework and then within the augmented Lagrangian framework. The fundamental idea of this algorithm is to search the solution of the contact problem no longer as the solution of a minimization problem but as the solution of the Euler equations of the minimization problem. The equation is then solved using a Newton method.

In addition, finding a good stopping criterion remains a difficulty. Very often, algorithms are stopped when the norm of the increment becomes sufficiently small. The main drawback of this technique is that there is no guarantee on the quality of the numerical solutions. Moreover, the solutions obtained by various algorithms are then difficult to compare. In this article, I propose a stopping criterion that can be used with the whole methods presented. The main advantage of this criterion lies in the fact that the mechanical quality of the solution is ensured (penetration, Signorini-Coulomb contact law, and equation of motion), both locally at each contact point and globally on the overall behavior.

The article is organized as follows: in the next part, I present the equations to be solved within the discrete element method context, and the frictional contact law considered. In the third part, I first present two classical methods to numerically solve the full problem, the first one based on the bipotential theory and the second one on the augmented Lagrangian theory. Then, I show how these methods can be enhanced using an appropriate Newton method, and a stopping criterion that works for all these algorithms is provided. The last part of this article is devoted to the numerical experiments to show the main properties of these algorithms.

## 2. PROBLEM SETTING

### 2.1. The equations of motion of a multicontact system

Classically (see, e.g., [6-9]), the motion of a multicontact system is described using a global generalized coordinate $\mathbf{q}$ (for $N_{p}$ particles, $\mathbf{q} \in \mathbb{R}^{\tilde{d} \times N_{p}}$, where $\tilde{d}=6$ for a three-dimensional (3D) problem and $\tilde{d}=3$ for a two-dimensional (2D) problem. Because of the possible shocks between particles, the equations of motion have to be formulated in terms of differential measure equation:

$$
\begin{equation*}
\mathbb{M} d \dot{\mathbf{q}}+\mathbf{F}^{\mathrm{int}}(t, \mathbf{q}, \dot{\mathbf{q}}) d t=\mathbf{F}^{\mathrm{ext}}(t, \mathbf{q}, \dot{\mathbf{q}}) d t+d R \tag{1}
\end{equation*}
$$

where

- $\mathbb{M}$ represents the generalized mass matrix;
- $\mathbf{F}^{\text {int }}$ and $\mathbf{F}^{\text {ext }}$ represent the internal and external forces, respectively;
- $d \mathbf{R}$ is a nonnegative real measure, representing the reaction forces and impulses between particles in contact.

For the sake of simplicity and without loss of generality, only the external forces are considered in the following. The internal forces are neglected because the general case can be easily derived through a linearizing procedure.

Then, for the numerics, Equation (1) is integrated on each time interval $\left[t_{k}, t_{k+1}\right]$ and approximated using a $\theta$-method with $\left.\theta \in] \frac{1}{2}, 1\right]$ for stability reasons (see $[10,11]$ ).

Therefore, the classical approximation of Equation (1) yields

$$
\left\{\begin{align*}
\mathbb{M}\left(\dot{\mathbf{q}}_{k+1}-\dot{\mathbf{q}}_{n}\right) & =\Delta t\left(\theta \mathbf{F}_{k+1}+(1-\theta) \mathbf{F}_{k}\right)+\mathbf{R}_{k+1}  \tag{2}\\
\mathbf{q}_{k+1} & =\mathbf{q}_{k}+\Delta t \theta \dot{\mathbf{q}}_{k+1}+\Delta t(1-\theta) \dot{\mathbf{q}}_{k}
\end{align*}\right.
$$

We will denote $\dot{\mathbf{q}}_{k}^{\text {free }}=\dot{\mathbf{q}}_{k}+\mathbb{M}^{-1} \Delta t\left(\theta \mathbf{F}_{k+1}+(1-\theta) \mathbf{F}_{k}\right.$ ) the free velocity (velocity when the contact forces vanish). Then, the first equation in (2) becomes

$$
\begin{equation*}
\dot{\mathbf{q}}_{k+1}=\dot{\mathbf{q}}_{k}^{\text {free }}+\mathbb{M}^{-1} \mathbf{R}_{k+1} \tag{3}
\end{equation*}
$$

To write the contact law, for a contact $c$ between two particles $\left(1 \leqslant c \leqslant N_{c}\right.$, where $N_{c}$ is the total number of contact), we define the local-global mapping

$$
\left\{\begin{array}{l}
\mathbf{u}^{c}=P^{*}(\mathbf{q}, c) \dot{\mathbf{q}}  \tag{4}\\
\mathbf{R}=P(\mathbf{q}, c) \mathbf{r}^{c}
\end{array}\right.
$$

where $\mathbf{u}^{c}$ is the local relative velocity between the two bodies in contact and $\mathbf{r}^{c}$ is the local contact force $\left(\mathbf{u}^{c}, \mathbf{r}^{c} \in \mathbb{R}^{d}\right.$ where $d$ is the dimension of the problem, and $P^{*}$ is the transpose of matrix $\left.P\right)$. We also denote $\mathbb{P}(\mathbf{q})$ the total-global mapping, for $\mathbf{u}$ and $\mathbf{r}$ in $\mathbb{R}^{d \times N_{c}}$ (vectors composed of all relative velocities and contact forces, respectively):

$$
\left\{\begin{array}{l}
\mathbf{u}=\mathbb{P}^{*}(\mathbf{q}) \dot{\mathbf{q}}  \tag{5}\\
\mathbf{R}=\mathbb{P}(\mathbf{q}) \mathbf{r}
\end{array}\right.
$$

In the discretization, a prediction of $\mathbf{q}$ is computed to estimate the mapping $\mathbb{P}(\mathbf{q})$ (see Equations (18) and (19) in the following).

Using Equations (2) and (5), we can write the discretization of the motion of a multicontact system, with frictional contact between particles, as

$$
\left\{\begin{align*}
\tilde{\mathbf{u}}_{k+1} & =\tilde{\mathbf{u}}_{k}^{\text {free }}+\mathbb{W} \mathbf{r}_{k+1}  \tag{6}\\
\operatorname{law}_{c}\left(\tilde{\mathbf{u}}_{k+1}^{c}, \mathbf{r}_{k+1}^{c}\right) & =\text { true } \quad \forall c \in\left\{1,2, \ldots, N_{c}\right\}
\end{align*}\right.
$$

where $\mathbb{W}=\mathbb{P}^{*} \mathbb{M}^{-1} \mathbb{P}$ is the Delassus operator, and $\tilde{\mathbf{u}}_{k}^{\text {free }}=\mathbb{P}^{*} \dot{\mathbf{q}}_{k}^{\text {free }}$ is the relative free velocity. Notice that a Newton impact law is also considered (see [8] and Equation (20) in the following), which modify $\mathbf{u}_{k}$ and $\mathbf{u}_{k}^{\text {free }}$ by $\tilde{\mathbf{u}}_{k}$ and $\tilde{\mathbf{u}}_{k}^{\text {free }}$, respectively.

The second equation in (6) is the implicit frictional contact law that is in our case the classical Signorini condition and Coulomb's friction law.

### 2.2. The frictional contact law

In the local coordinates system defined by the local normal vector $\mathbf{n}$ and the tangential vector $\mathbf{t} \perp \mathbf{n}$, any element $\mathbf{u}$ and $\mathbf{r}$ can be uniquely decomposed as $\mathbf{u}=u_{n} \mathbf{n}+\mathbf{u}_{t}$ and $\mathbf{r}=r_{n} \mathbf{n}+\mathbf{r}_{t}$, respectively. Using these coordinates, we can state the unilateral contact law using the Signorini's conditions (see Figure 1 for a graphical representation):

$$
\begin{equation*}
u_{n} \geqslant 0, \quad r_{n} \geqslant 0, \quad u_{n} r_{n}=0 \tag{7}
\end{equation*}
$$

On the other hand, the Coulomb's law of friction can be stated using the algorithmic form (see Figure 2 for a graphical representation):
$\left[\begin{array}{lll}\text { If } r_{n}=0 & \text { then } u_{n} \geqslant 0 & \text { ! No contact } \\ \text { Else if } r_{n}>0 \text { and }\left\|\mathbf{r}_{t}\right\|<\mu r_{n} & \text { then } \mathbf{u}=0 & \text { ! Sticking } \\ \text { Else } r_{n}>0 \text { and }\left\|\mathbf{r}_{t}\right\|=\mu r_{n} & \text { then } \exists \lambda \geqslant 0 \text { such that } \mathbf{u}_{t}=\lambda \frac{\mathbf{r}_{t}}{\left\|r_{t}\right\|} & \text { ! Sliding }\end{array}\right.$

Else $r_{n}>0$ and $\left\|\mathbf{r}_{t}\right\|=\mu r_{n} \quad$ then $\exists \lambda \geqslant 0$ such that $\mathbf{u}_{t}=\lambda \frac{\mathbf{r}_{t}}{\left\|\mathbf{r}_{t}\right\|} \quad$ ! Sliding


Figure 1. The Signorini conditions.


Figure 2. The Coulomb's conditions.


Figure 3. The Coulomb's cone.

For a given friction coefficient $\mu$, let $K_{\mu}$ be the isotropic Coulomb's cone, which defines the set of admissible forces (Figure 3):

$$
\begin{equation*}
K_{\mu}=\left\{\mathbf{r}=r_{n} \mathbf{n}+\mathbf{r}_{t}:\left\|\mathbf{r}_{t}\right\|-\mu r_{n} \leqslant 0\right\} \tag{9}
\end{equation*}
$$

The previous law can also be written as follows:
$\left[\begin{array}{lll}\text { If } r_{n}=0 & \text { then } u_{n} \geqslant 0 & \text { ! No contact } \\ \text { Else if } \mathbf{r} \in I\left(K_{\mu}\right) & \text { then } \mathbf{u}=0 & \text { ! Sticking } \\ \text { Else } r_{n}>0 \text { and } \mathbf{r} \in B\left(K_{\mu}\right) & \text { then } \exists \lambda \geqslant 0 \text { such that } \mathbf{u}_{t}=\lambda \frac{\mathbf{r}_{t}}{\left\|\mathbf{r}_{t}\right\|} & !\text { Sliding }\end{array}\right.$
where $I\left(K_{\mu}\right)$ and $B\left(K_{\mu}\right)$ are, respectively, the interior and the boundary of the cone $K_{\mu}$.

## 3. NUMERICAL RESOLUTION OF THE CONTACT/FRICTION PROBLEMS

I will describe in this section the numerical algorithms that will be considered in what follows. Generally, to solve problem (6), the numerical algorithms considered are based on two levels: the global level where the equations of motion are solved and the local level devoted to the resolution of the contact law.

### 3.1. Resolution of the global problem : the nonlinear Gauss-Seidel (NLGS) method

In this paragraph, I describe the algorithm used at the global level to solve problem (6). Following the ideas of Jean and Moreau [7,8, 12], I use the NLGS algorithm, which is the most commonly used. It consists in considering successively each contact until the convergence. The numerical criterion used to state the convergence will be studied later in this article.

This method is intrinsically sequential, but it is possible to use a simple multithreading technique that consists in splitting the contact loop onto several threads. This method has been studied in [13] in the case of a local algorithm based on the augmented Lagrangian method.

Notice that it is also possible to consider at this stage more sophisticated methods such as conjugate gradient-type methods (see, e.g., [11]). However, these techniques do not improve significantly the speedup of the convergence. That is why, they are not considered in the following.

### 3.2. The standard bipotential-based method (SBP)

In this paragraph, I provide a first method to solve the contact problem at the local level (contact point between two particles). The method is based on the notion of bipotential, introduced by de Saxcé et al. [1].

Using the bipotential framework, we can easily show (see, e.g., [1, 14-16]) that a couple (u,r) verifies the Signorini-Coulomb contact rules if and only if

$$
\begin{equation*}
b_{c}(\mathbf{v}, \mathbf{s})+\mathbf{v} \cdot \mathbf{s} \geqslant b_{c}(\mathbf{u}, \mathbf{r})+\mathbf{u} \cdot \mathbf{r}=0 \quad \forall \mathbf{v}, \mathbf{s} \tag{11}
\end{equation*}
$$

where $b_{c}$ is the bipotential

$$
\begin{equation*}
b_{c}(-\mathbf{u}, \mathbf{r})=\Psi_{\mathbb{R}^{+}}\left(u_{n}\right)+\Psi_{K_{\mu}}(\mathbf{r})+\mu r_{n}\left\|u_{t}\right\| \tag{12}
\end{equation*}
$$

and $\Psi_{C}$ stands for the indicator function of the set $C: \Psi_{C}(x)=0$ if $x \in C, \Psi_{C}(x)=+\infty$ if $x \notin C$.
Consequently, the contact law can be written in a compact form of an implicit subnormality rule (or a differential inclusion rule):

$$
\begin{equation*}
-\mathbf{u} \in \partial_{\mathbf{r}} b_{c}(-\mathbf{u}, \mathbf{r}) \tag{13}
\end{equation*}
$$

Then, for a contact $c$, at an NLGS iteration $i$, knowing the relative velocity $\tilde{\mathbf{u}}^{c, i}$, the algorithm to compute $\mathbf{r}^{c, i+1}$ from $\mathbf{r}^{c, i}$ is based on the minimization of the bipotential (see, e.g.,[14]), using the inequality

$$
\begin{equation*}
b_{c}\left(-\tilde{\mathbf{u}}^{c, i}, \mathbf{r}\right)+\tilde{\mathbf{u}}^{c, i} \cdot \mathbf{r} \geqslant b_{c}\left(-\tilde{\mathbf{u}}^{c, i}, \mathbf{r}^{c, i+1}\right)+\tilde{\mathbf{u}}^{k, i} \cdot \mathbf{r}^{c, i+1} \quad \forall \mathbf{r} \in K_{\mu} \tag{14}
\end{equation*}
$$

or $g(\mathbf{r}) \geqslant g\left(\mathbf{r}^{c, i+1}\right), \forall \mathbf{r} \in K_{\mu}$, if we denote

$$
\begin{equation*}
g(\mathbf{r})=\Psi_{\mathbb{R}^{+}}\left(\tilde{u}_{n}^{c, i}\right)+\Psi_{K_{\mu}}(\mathbf{r})+\mu r_{n}\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|+\tilde{\mathbf{u}}^{c, i} \cdot \mathbf{r} \tag{15}
\end{equation*}
$$

The minimization of (14) is classically realized using a projected gradient method (Uzawa method) without considering the singular term $\Psi_{\mathbb{R}^{+}}\left(\tilde{u}_{n}^{c, i}\right)$. This minimization can also be viewed as the search of the proximal point of the augmented force $\mathbf{r}-\rho \tilde{\mathbf{u}}$, with respect to the function $\mathbf{r} \mapsto \rho b_{c}(-\tilde{\mathbf{u}}, \mathbf{r})$ (see, e.g., $\left.[1,14,15]\right):$

$$
\mathbf{r}=\operatorname{prox}\left(\mathbf{r}-\rho \tilde{\mathbf{u}}, \rho b_{c}(-\tilde{\mathbf{u}}, \mathbf{r})\right)
$$

More precisely, the Uzawa method leads to compute the augmented force $\tau^{c, i+1}=\mathbf{r}^{c, i}-$ $\rho \nabla \tilde{g}\left(\mathbf{r}^{c, i}\right)$, where $\tilde{g}$ is the differential part of $g$ :

$$
\nabla \tilde{g}\left(\mathbf{r}^{c, i}\right)=\nabla_{\mathbf{r}}\left(\mu r_{n}\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|+\tilde{\mathbf{u}}^{c, i} \cdot \mathbf{r}\right)=\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\| \mathbf{n}+\tilde{\mathbf{u}}^{c, i}
$$

and to consider the force at the next step as a projection of the augmented force onto the set of admissible force $\mathbf{r}^{c, i+1}=\operatorname{proj}\left(\tau^{c, i+1}, K_{\mu}\right)$, which provides Equations (21) and (22) in the resolution algorithm of the global problem presented in the following paragraphs. The $\operatorname{proj}\left(\tau^{c, i+1}, K_{\mu}\right)$ stands for the orthogonal projection over the convex $K_{\mu}$, which can be computed exactly (see [14]).

This algorithm will be referred as the SBP method thereafter.
For the sake of simplicity, we denote hereafter the descent direction

$$
\mathbf{D}^{c, i}=\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\| \mathbf{n}+\tilde{\mathbf{u}}^{c, i}
$$

## Remark 1

A first improvement of this method could be to compute the optimal step $\rho^{c, i}$. To do so, we have to minimize

$$
\begin{equation*}
\rho \mapsto g\left(\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i}\right) \tag{16}
\end{equation*}
$$

or, more precisely,

$$
\begin{align*}
\rho & \mapsto \Psi_{\mathbb{R}^{+}}\left(\tilde{u}_{n}^{c, i}\right)+\Psi_{K_{\mu}}\left(\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i}\right)+\mu\left(r_{n}^{c, i}-\rho \mathbf{D}^{c, i} \cdot \mathbf{n}\right)\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|+\tilde{\mathbf{u}}^{c, i} \cdot\left(\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i}\right) \\
& =\Psi_{\mathbb{R}^{+}}\left(\tilde{u}_{n}^{c, i}\right)+\Psi_{K_{\mu}}\left(\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i}\right)-\rho \mathbf{D}^{c, i} \cdot\left(\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\| \mathbf{n}+\tilde{\mathbf{u}}^{c, i}\right)+C t e \\
& =\Psi_{\mathbb{R}^{+}}\left(\tilde{u}_{n}^{c, i}\right)+\Psi_{K_{\mu}}\left(\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i}\right)-\rho\left\|\mathbf{D}^{c, i}\right\|^{2}+C t e \tag{17}
\end{align*}
$$

We can observe that this method does not permit to choose an optimal parameter $\rho$ because $g$, as a function of $\rho$, is linear, excepted in the case where it exists $\rho>0$ such that $\mathbf{r}^{c, i}-\rho \mathbf{D}^{c, i} \notin K_{\mu}$. A solution could be to modify the function $g$, for example, by replacing $\tilde{\mathbf{u}}^{c, i}$ by a prediction of $\tilde{\mathbf{u}}^{c, i+1}$ using the equations of the dynamics. Unfortunately, this method does not provide satisfying numerical results.

Then, the SBP algorithm can be written (see, e.g., [16]):

- Loop on the step time $k$
- Prediction of a position (for the computation of the local-global mapping):

$$
\begin{equation*}
\mathbf{q}_{k+\frac{1}{2}}=\mathbf{q}_{k}+\frac{\Delta t}{2} \dot{\mathbf{q}}_{k} \tag{18}
\end{equation*}
$$

- Initialization of the motion: $\dot{\mathbf{q}}_{k+1}^{0}=\dot{\mathbf{q}}_{k}^{\text {free }}$ (initialization of the contact forces with $\mathbf{R}=0$ ).
- Loop on $i \geqslant 0$ (NLGS), until convergence
* Loop on the contacts $c$ :
- Computation of the local-global mapping

$$
\begin{equation*}
\dot{\mathbf{u}}^{-}=P^{*}\left(\mathbf{q}_{k+\frac{1}{2}}, c\right) \dot{\mathbf{q}}_{k} ; \quad \dot{\mathbf{u}}^{c,+i}=P^{t}\left(\mathbf{q}_{k+\frac{1}{2}}, c\right) \dot{\mathbf{q}}_{k+1}^{i} \tag{19}
\end{equation*}
$$

- Newton shock law

$$
\begin{equation*}
\tilde{u}_{n}^{c, i}=\frac{u_{n}^{c,+i}+e_{n} u_{n}^{-}}{1+e_{n}} ; \quad \tilde{\mathbf{u}}_{t}^{c, i}=\frac{\mathbf{u}_{t}^{c,+i}+e_{n} \mathbf{u}_{t}^{-}}{1+e_{t}} \tag{20}
\end{equation*}
$$

- Prediction of the reaction:

$$
\begin{equation*}
\tau^{c, i+1}=\mathbf{r}^{c, i}-\rho\left[\tilde{\mathbf{u}}_{t}^{c, i}+\left(\tilde{u}_{n}^{c, i}+\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|\right) \mathbf{n}\right] \tag{21}
\end{equation*}
$$

- Correction of the reaction:

$$
\begin{equation*}
\mathbf{r}^{c, i+1}=\operatorname{proj}\left(\tau^{c, i+1}, K_{\mu}\right) \tag{22}
\end{equation*}
$$

- Actualization of the generalized displacement:

$$
\begin{equation*}
\dot{\mathbf{q}}_{k+1}^{i+1}=\dot{\mathbf{q}}_{k}^{\mathrm{frre}}+\mathbb{M}^{-1}\left(\sum_{\alpha \leqslant c} P\left(\mathbf{q}_{k+\frac{1}{2}}, \alpha\right) \mathbf{r}^{\alpha, i+1}+\sum_{\alpha>c} P\left(\mathbf{q}_{k+\frac{1}{2}}, \alpha\right) \mathbf{r}^{\alpha, i}\right) \tag{23}
\end{equation*}
$$

* End of the loop on contacts $c$.
- End of the loop on $i$ of NLGS. When the convergence is reached, actualization of the velocity $\dot{\mathbf{q}}_{k+1}=\dot{\mathbf{q}}_{k+1}^{i+1}$
- Actualization of the generalized displacements: $\mathbf{q}_{k+1}=\mathbf{q}_{k+\frac{1}{2}}+\frac{\Delta t}{2} \dot{\mathbf{q}}_{k+1}$
- End of the loop on the step time $k$.


## Remark 2

Notice that only one iteration of the Uzawa algorithm at the local level is considered. Various previous studies (see, e.g., [17]) show that there is no significant improvement of the method if several iterations of the Uzawa algorithm are considered at this stage.

### 3.3. Newton method and enhanced bipotential method (EBP)

I introduce in this section a Newton method to speed up the convergence in the computation of the solution. This method has been already used, especially in the case of the augmented Lagrangian method developed by Alart et al. [2], and the ideas presented in this article follow those of Feng et al. [3] and have been adapted to the problem of the discrete element method. The main idea of this technique is to find the solution of the optimization problem, not as a minimum but as a zero of a function, using the Euler equation (or the saddle point) of the problem. Then, a standard Newton method can be developed to solve this Euler equation.

The technique is first described in the case of the bipotential framework and will be adapted to the augmented Lagrangian method farther.

We recall that the local problem to be solved, for each contact $c$, can be written as

$$
\left\{\begin{array}{l}
\tilde{\mathbf{u}}_{k+1}^{c}=\tilde{\mathbf{u}}_{k}^{c, \text { free }}+\sum_{\alpha=1}^{N_{c}} W_{c \alpha} \mathbf{r}^{\alpha}  \tag{24}\\
\mathbf{r}^{c}=\operatorname{proj}\left(\tau^{c}, K_{\mu}\right)
\end{array} \quad \forall c=1, \ldots, N_{c}\right.
$$

where $\tau^{c}=\mathbf{r}^{c}-\rho\left(\mu\left\|\tilde{\mathbf{u}}_{t}^{c}\right\| \mathbf{n}+\tilde{\mathbf{u}}\right)$ is the augmented reaction (see (21)), and $W_{c \alpha}=$ $P^{*}\left(\mathbf{q}_{k+\frac{1}{2}, c}\right) \mathbb{M}^{-1} P\left(\mathbf{q}_{k+\frac{1}{2}, \alpha}\right)$ is the local Delassus operator.

This problem can be written equivalently

$$
\left\{\begin{array}{l}
\tilde{\mathbf{u}}_{k+1}^{c}-\tilde{\mathbf{u}}_{k}^{c, \text { free }}-\sum_{\alpha=1}^{N_{c}} W_{c \alpha} \mathbf{r}^{\alpha}=0  \tag{25}\\
\mathbf{r}^{c}-\operatorname{proj}\left(\tau^{c}, K_{\mu}\right)=0
\end{array} \quad \forall c=1, \ldots, N_{c}\right.
$$

Reminding now that we want to use a Newton algorithm to solve theses equations inside the NLGS loop on the variable $i$, we define now, for each contact $c=1, \ldots, N_{c}$, the function

$$
f_{c}^{i}(\chi)=\binom{\tilde{\mathbf{u}}^{c, i}-\tilde{\mathbf{u}}_{k}^{c, \text { free }}-\sum_{\alpha=1}^{N_{c}} W_{c \alpha} \mathbf{r}^{\alpha, i}}{\mathbf{Z}^{c, i}}
$$

where

- the vector $\mathbf{Z}^{c}$ is the error on the prediction of the reaction

$$
\begin{equation*}
\mathbf{Z}^{c, i}\left(\mathbf{r}^{c, i}, \tilde{\mathbf{u}}^{c, i}\right)=\mathbf{r}^{c, i}-\operatorname{proj}\left(\tau^{c, i}, K_{\mu}\right) \tag{26}
\end{equation*}
$$

- $\chi_{c}=\left(\mathbf{r}^{c, i}, \tilde{\mathbf{u}}^{c, i}\right)^{t}$
- $\chi=\left(\chi_{1}, \chi_{2}, \ldots, \chi_{N_{c}}\right)^{t}$


## Remark 3

The first equality in the relation $f(\chi)=0$ is the equation of motion for the bodies in contact, and the second relation is the frictional Coulomb law between the bodies in contact, written within the bipotential framework.

Then, we have to write a Newton algorithm to solve the problem $f(\chi)=0$. This algorithm can be written, for a contact $c$, by substituting Equations (21) and (22) in algorithm (SBP) by the following:

- Initialization:

$$
\chi_{c}^{0}=\left(\mathbf{r}^{0}=\mathbf{r}^{c, i}, \mathbf{v}^{0}=\tilde{\mathbf{u}}^{c, i}\right)^{t}, \quad \ell=0
$$

- Loop on $\ell$, until convergence:
$-\tau_{\ell}^{c}=\mathbf{r}^{\ell}-\rho\left(\mu\left\|\mathbf{v}_{t}^{\ell}\right\| \mathbf{n}+\mathbf{v}^{\ell}\right)$
- Resolution:

$$
\begin{equation*}
\left[\frac{\partial f_{c}}{\partial \chi^{c}}\left(\chi^{\ell}\right)\right] \Delta \chi_{c}=-f_{c}\left(\chi^{\ell}\right) \tag{27}
\end{equation*}
$$

- Actualization: $\chi_{c}^{\ell+1}=\chi_{c}^{\ell}+\Delta \chi_{c}$
- End of the loop on $\ell$ until convergence, $\tilde{\mathbf{u}}^{c, i+1}=\mathbf{v}^{\ell}$ and $\mathbf{r}^{c, i+1}=\mathbf{r}^{\ell}$.


## Remark 4

This algorithm needs more than one iteration at each NLGS iteration to be efficient. As a consequence and compared with the Uzawa algorithm, the solution in the Newton algorithm is controlled by both the local (iteration $\ell$ ) and global convergence criteria (iteration $i$, see [3, 17]).

The local convergence criterion for the Newton algorithm is defined by

$$
\begin{equation*}
\varepsilon_{\text {Newt }}^{c}\left(\chi_{\ell}\right)=\left\|\mathbf{v}^{\ell}-\mathbf{u}_{k}^{c \text { ffree }}-W \mathbf{r}^{\ell}\right\|+\left\|\mathbf{r}^{\ell}-\operatorname{proj}\left(\mathbf{r}^{\ell}, K_{\mu}\right)\right\| \tag{28}
\end{equation*}
$$

This criterion measures $f_{c}\left(\chi_{\ell}\right)$ that has to be sufficiently small.
The matrix $\left[\partial f_{c} / \partial \chi_{c}(\chi)\right]$ represents the tangential matrix of the local equations for the contact $c$. This matrix is of dimension $6 \times 6$ for a 3 D problem and $4 \times 4$ for a 2D problem. For a 3 D problem, the general form of this matrix is the following:

$$
\left[\frac{\partial f_{c}}{\partial \chi_{c}}(\chi)\right]=\left[\begin{array}{cc}
-W & I d_{3 \times 3}  \tag{29}\\
A_{c} & B_{c}
\end{array}\right]
$$

where

$$
\begin{equation*}
A_{c}=\left[\frac{\partial Z_{c}}{\partial r_{n}}\left|\frac{\partial Z_{c}}{\partial r_{t_{1}}}\right| \frac{\partial Z_{c}}{\partial r_{t_{2}}}\right] \quad B_{c}=\left[\frac{\partial Z_{c}}{\partial v_{n}}\left|\frac{\partial Z_{c}}{\partial v_{t_{1}}}\right| \frac{\partial Z_{c}}{\partial v_{t_{2}}}\right] \tag{30}
\end{equation*}
$$

The matrices $A_{c}$ and $B_{c}$ take different forms according to the contact status:

- First case: sliding contact.

In that case, we have

$$
\mu\left\|\tau_{t}\right\| \geqslant-\tau_{n} \quad\left\|\tau_{t}\right\| \geqslant \mu \tau_{n}
$$

then

$$
\operatorname{Proj}\left(\tau, K_{\mu}\right)=\tau-\left(\frac{\left\|\tau_{t}\right\|-\mu \tau_{n}}{1+\mu^{2}}\right)\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)
$$

and

$$
\mathbf{Z}_{c}=\rho\left(\mu\left\|\mathbf{v}_{t}^{k}\right\| \mathbf{n}+\mathbf{v}^{k}\right)+\left(\frac{\left\|\tau_{t}\right\|-\mu \tau_{n}}{1+\mu^{2}}\right)\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)
$$

The computation of the derivatives of $\mathbf{Z}_{c}$ provides the matrices $A_{c}$ and $B_{c}$ :

$$
\begin{aligned}
-\frac{\partial \mathbf{Z}_{c}}{\partial r_{n}} & =-\frac{\mu}{1+\mu^{2}}\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial r_{t_{1}}} & =\frac{\tau_{t_{1}}}{\left(1+\mu^{2}\right)\left\|\tau_{t}\right\|}\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)+\frac{\left\|\tau_{t}\right\|-\mu \tau_{n}}{1+\mu^{2}}\left(\frac{\mathbf{t}_{1}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{1}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial r_{t_{2}}} & =\frac{\tau_{t_{2}}}{\left(1+\mu^{2}\right)\left\|\tau_{t}\right\|}\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)+\frac{\left\|\tau_{t}\right\|-\mu \tau_{n}}{1+\mu^{2}}\left(\frac{\mathbf{t}_{2}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{2}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{n}} & =\rho \mathbf{n}+\frac{\rho \mu}{1+\mu^{2}}\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t_{1}}} & =\rho\left(\mathbf{t}_{1}+\mu \frac{v_{t_{1}}}{\left\|\mathbf{v}_{t}\right\|} \mathbf{n}\right)-\frac{\rho}{1+\mu^{2}}\binom{\left(\frac{\tau_{t_{1}}}{\left\|\tau_{t}\right\|}-\frac{\mu^{2} v_{t_{1}}}{\left\|v_{t}\right\|}\right)\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)+}{\left(\left\|\tau_{t}\right\|-\mu \tau_{n}\right)\left(\frac{\mathbf{t}_{1}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{1}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right)} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t_{2}}} & =\rho\left(\mathbf{t}_{2}+\mu \frac{v_{t_{2}}}{\left\|\mathbf{v}_{t}\right\|} \mathbf{n}\right)-\frac{\rho}{1+\mu^{2}}\binom{\left(\frac{\tau_{t_{2}}}{\left\|\tau_{t}\right\|}-\frac{\mu^{2} v_{t_{2}}}{\left\|\mathbf{v}_{t}\right\|}\right)\left(\frac{\tau_{t}}{\left\|\tau_{t}\right\|}-\mu \mathbf{n}\right)+}{\left(\left\|\tau_{t}\right\|-\mu \tau_{n}\right)\left(\frac{\mathbf{t}_{2}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{2}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right)}
\end{aligned}
$$

For a 2D problem, these computations yield the following:

$$
\begin{aligned}
-\frac{\partial \mathbf{Z}_{c}}{\partial r_{n}} & =\frac{\mu}{1+\mu^{2}}\left(\mu \mathbf{n}-\theta_{r} \mathbf{t}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial r_{t}} & =\frac{1}{\left(1+\mu^{2}\right)}\left(-\mu \theta_{r} \mathbf{n}+\mathbf{t}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{n}} & =\frac{\rho}{1+\mu^{2}}\left(\mathbf{n}+\mu \theta_{r} \mathbf{t}\right) \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t}} & =\frac{\rho \mu}{1+\mu^{2}}\left(\left(\theta_{v}+\theta_{r}\right) \mathbf{n}+\mu\left(1-\theta_{r} \theta_{v}\right) \mathbf{t}\right)
\end{aligned}
$$

where $\theta_{v}=\operatorname{sign}\left(\mathbf{v}_{t}\right)$ and $\theta_{r}=\operatorname{sign}\left(\tau_{t}\right)$.

- Second case: sticking contact.

In that case, we have

$$
\mu\left\|\tau_{t}\right\| \geqslant-\tau_{n} \quad\left\|\tau_{t}\right\|<\mu \tau_{n}
$$

then

$$
\mathbf{Z}_{c}=\rho\left(\mu\left\|\mathbf{v}_{t}^{k}\right\| \mathbf{n}+\mathbf{v}^{k}\right)
$$

and the computation of the derivatives of $\mathbf{Z}_{c}$ reads

$$
\begin{aligned}
-A_{c} & =0_{3 \times 3} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{n}} & =\rho \mathbf{n} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t_{1}}} & =\rho \mu \frac{v_{t_{1}}}{\left\|\mathbf{v}_{t}\right\|} \mathbf{n}+\rho \mathbf{t}_{1} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t_{2}}} & =\rho \mu \frac{v_{t_{2}}}{\left\|\mathbf{v}_{t}\right\|} \mathbf{n}+\rho \mathbf{t}_{2}
\end{aligned}
$$

For a 2D problem, these computations lead to the following:

$$
\begin{aligned}
-A_{c} & =0_{2 \times 2} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{n}} & =\rho \mathbf{n} \\
-\frac{\partial \mathbf{Z}_{c}}{\partial v_{t}} & =\rho \mu \theta_{v} \mathbf{n}+\rho \mathbf{t}
\end{aligned}
$$

- Third case: no contact.

In that case, the matrices $A_{c}=\operatorname{Id}_{3 \times 3}$ and $B_{c}$ vanish, and $\chi_{c}^{\ell+1}=\left\{\begin{array}{c}0 \\ \mathbf{v}^{k}\end{array}\right\}$

### 3.4. Resolution of the linear system

Generally, the drawback of a Newton method is the computational cost of the linear system to be solved at each iteration. Here, the particular form of the tangent matrix allows the use of a condensation technique. More precisely, the linear system to be solved can be written as

$$
\left[\begin{array}{cc}
-W & I d_{3 \times 3}  \tag{31}\\
A_{c} & B_{c}
\end{array}\right]\binom{\delta \mathbf{r}}{\delta \mathbf{v}}=\binom{-\mathbf{f}}{-\mathbf{g}}
$$

The first equation yields $\delta \mathbf{v}=-\mathbf{f}+W \delta \mathbf{r}$, and introducing this equality in the second equation leads to solve the linear system

$$
\begin{equation*}
\left(A_{c}+B_{c} W\right) \delta \mathbf{r}=-\mathbf{g}+B_{c} \mathbf{f} \tag{32}
\end{equation*}
$$

This property halves the size of the linear system to be solved.

## Remark 5

A drawback of the bipotential framework is that, due to its specificity, it is rather difficult to consider fully coupled problems, where the contact law and another phenomena, such as electricity or thermal effects, are strongly coupled. The other method presented in this article has a better property from this point of view because it is based on a more standard mathematical background in the theory of optimization (see, e.g., [18]).

### 3.5. Standard augmented Lagrangian (SAL) and enhanced augmented Lagrangian (EAL) method

In [2], Alart et al. proposed another method to solve the frictional contact problem. This method has been also used with various improvement (e.g., parallelization and conjugate gradient method) to solve multicontact problems (see, e.g., [11, 13, 19-22]). Even if the coupled frictional contact problem is not an optimization problem anymore, it is always possible to formally formulate a 'quasi'-optimization problem, for which the constraint set depends on the normal components of the solution as a parameter. The solution is then searched as a saddle point of a quasi-augmented Lagrangian of the problem.

More precisely, the global problem on all unknowns that has to be solved at each time step (in place of Equation (24)) has the following form:

$$
\left\{\begin{array}{l}
\mathbf{u}=\mathbf{u}^{\text {free }}+\mathbb{W} \mathbf{r}  \tag{33}\\
\mathbf{r} \geqslant 0, \mathbf{u} \geqslant 0, \mathbf{r} \cdot \mathbf{u}=0
\end{array}\right.
$$

To solve this problem, for a given $\mathbf{r} \in \mathbb{R}^{3 \times N_{c}}$, one can define the Cartesian product of infinite half-cylinder with section equal to the ball $\mathcal{B}\left(0, \mu r_{c}\right)$ of radius $\mu r_{c}$ by:

$$
\mathcal{C}(\mu \mathbf{r})=\prod_{c=1}^{N_{c}} \mathbb{R}^{+} \times \mathcal{B}\left(0, \mu r_{c}\right)
$$

and then, the granular-type frictional contact problem is given by

$$
\begin{equation*}
\mathbf{r} \in \operatorname{argmin}_{\mathbf{r} \in \mathcal{C}(\mu \mathbf{r})} \frac{1}{2} \mathbf{r} \cdot \mathbb{W} \mathbf{r}+\mathbf{u}^{\text {free }} \cdot \mathbf{r}=\operatorname{argmin}_{\mathbf{r} \in \mathcal{C}(\mu \mathbf{r})} J(\mathbf{r}) \tag{34}
\end{equation*}
$$

and the projected gradient method to minimize this problem reads (for each iteration $i$ of the NLGS algorithm) as follows:

$$
\begin{equation*}
\mathbf{r}^{i+1}=\operatorname{proj}\left(\mathbf{r}^{i}-\rho\left(\mathbf{u}^{\text {free }}+\mathbb{W} \mathbf{r}^{i}\right), \mathcal{C}\left(\mu \mathbf{r}^{i+1}\right)\right) \tag{35}
\end{equation*}
$$

or $\mathbf{r}^{i+1}=\operatorname{proj}\left(\tau^{i+1}, \mathcal{C}\left(\mu \mathbf{r}^{i+1}\right)\right)$, with $\tau^{i+1}=\mathbf{r}^{i}-\rho \mathbf{u}^{i}, \mathbf{u}^{i}=\mathbf{u}^{\text {free }}+\mathbb{W} \mathbf{r}^{i}$. This algorithm will be referred to hereinafter as the SAL method.

Notice that this method is very close to the SBP method. More precisely, for a contact $c$, only the descent direction $\tilde{\mathbf{u}}^{c, i}+\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\| \mathbf{n}$ in (21) is replaced by $\tilde{\mathbf{u}}^{c, i}$, and the projection $\mathbf{r}^{c, i+1}=$ $\operatorname{proj}\left(\tau^{c, i+1}, K_{\mu}\right)$ in (22) is replaced by

$$
\left\{\begin{array}{l}
r_{n}^{c, i+1}=\max \left(0, \tau_{n}^{c, i+1}\right) \\
\mathbf{r}_{t}^{c, i+1}=\frac{\tau_{t}^{c, i+1}}{\left\|\tau_{t}^{c, i+1}\right\|} \min \left(\mu r_{n}^{c, i+1},\left\|\tau_{t}^{c, i+1}\right\|\right)
\end{array}\right.
$$

Remark 6
It is also possible to see the algorithm developed from the bipotential formalism as a slight modification of the algorithm defined previously. Indeed, it is only necessary to change the set $\mathcal{C}(\mathbf{r})$ by $\mathcal{K}=\prod_{c=1}^{N_{c}} K_{\mu}$, and to change the descent direction $\tilde{\mathbf{u}}^{c, i}$ by $\tilde{\mathbf{u}}^{c, i}+\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\| \mathbf{n}$, which remains a descent direction for the SAL method, because

$$
\nabla J\left(\mathbf{r}^{c, i+1}\right) \cdot \mathbf{D}^{c, i}=-\left\|\tilde{\mathbf{u}}^{c, i}\right\|^{2}-\tilde{\mathbf{u}}^{c, i} \cdot\left(\mu\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|\right) \mathbf{n}=-\left\|\tilde{\mathbf{u}}^{c, i}\right\|^{2}-\mu u_{n}^{c, i}\left\|\tilde{\mathbf{u}}_{t}^{c, i}\right\|
$$

which is negative because $\mu \geqslant 0$.

Then, acting by analogy, we can develop a Newton method to find the minimum of $J$ by seeking the solution as a zero of the function $\tilde{f}(\chi)$ where, for a contact $c$,

$$
\tilde{f}_{c}(\chi)=\binom{\tilde{\mathbf{u}}_{k+1}^{c}-\tilde{\mathbf{u}}_{k}^{c \text { free }}-\sum_{\alpha=1}^{N_{c}} W_{c \alpha} \mathbf{r}^{\alpha}}{\tilde{\mathbf{Z}}^{c}}
$$

the vector $\tilde{\mathbf{Z}}^{c}$ is the error on the prediction of the reaction

$$
\begin{equation*}
\tilde{\mathbf{Z}}^{c}\left(\mathbf{r}^{c}, \tilde{\mathbf{u}}_{k+1}^{c}\right)=\mathbf{r}^{c}-\operatorname{proj}\left(\tau_{k+1}^{c}, \mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)\right) \tag{36}
\end{equation*}
$$

and the set $\mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)$ is the set of admissible forces $\mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)=\mathbb{R}^{+} \times \mathcal{B}\left(0, r_{c}\right)$. This method will be referred as the EAL method hereafter.

Then, in what follows, we have three cases in the computation of the tangent matrix $\left[\partial \tilde{f} / \partial \chi^{c}\left(\chi^{\ell}\right)\right]:$

- First case: sliding contact $\left(\tau_{n}>0, \tau_{t} \geqslant \mu r_{n}\right)$

We have $\operatorname{proj}\left(\tau^{c}, \mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)\right)=\tau_{n} \mathbf{n}+\frac{\tau_{t}}{\left\|\tau_{t}\right\|} \mu r_{n}$ and $\tilde{\mathbf{Z}}_{c}=\rho v_{n} \mathbf{n}-\frac{\tau_{t}}{\left\|\tau_{t}\right\|} \mu r_{n}+\mathbf{r}_{t}$.

The computation of the derivatives of $\tilde{\mathbf{Z}}_{c}$ provides the matrices $A_{c}$ and $B_{c}$ :

$$
\begin{aligned}
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial r_{n}}=-\mu \frac{\tau_{t}}{\left\|\tau_{t}\right\|} \\
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial r_{t_{1}}}=\mathbf{t}_{1}-\mu r_{n}\left(\frac{\mathbf{t}_{1}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{1}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right) \\
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial r_{t_{2}}}=\mathbf{t}_{2}-\mu r_{n}\left(\frac{\mathbf{t}_{2}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{2}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right) \\
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial v_{n}}=\rho \mathbf{n} \\
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial v_{t_{1}}}=-\rho \mu r_{n}\left(\frac{\mathbf{t}_{1}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{1}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right) \\
& -\frac{\partial \tilde{\mathbf{Z}}_{c}}{\partial v_{t_{2}}}=-\rho \mu r_{n}\left(\frac{\mathbf{t}_{2}}{\left\|\tau_{t}\right\|}-\frac{\tau_{t_{2}}}{\left\|\tau_{t}\right\|^{3}} \tau_{t}\right)
\end{aligned}
$$

For a 2D problem, these computations lead to

$$
A_{c}=\left(\begin{array}{cc}
0 & 0 \\
-\mu \theta_{r} & 1
\end{array}\right) \quad B_{c}=\rho\left(\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right)
$$

- Second case: sticking contact ( $\tau_{n}>0, \tau_{t}<\mu r_{n}$ )
$\operatorname{proj}\left(\tau^{c}, \mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)\right)=\tau^{c}$, and the computation of the derivatives of $\mathbf{Z}_{c}$ reads

$$
\begin{aligned}
& -A_{c}=0_{3 \times 3} \\
& -B_{c}=\rho \mathrm{Id}_{3 \times 3}
\end{aligned}
$$

- Third case: no contact $\left(\tau_{n} \leqslant 0\right)$
$\operatorname{proj}\left(\tau^{c}, \mathcal{C}_{c}\left(\mu \mathbf{r}^{c}\right)\right)=0$, then the matrices $A_{c}=\operatorname{Id}_{3 \times 3}$ and $B_{c}$ vanishes, and $\chi_{c}^{\ell+1}=\left\{\begin{array}{c}0 \\ \mathbf{v}^{k}\end{array}\right\}$


### 3.6. The global stopping (convergence) criterion

I present in this paragraph the convergence criterion on the global NLGS iterations. This criterion, developed from that proposed in [23], has been extended in the case of the Newton and bipotential (EBP) method, where some terms are naturally vanishing in the original Uzawa and bipotential (SBP) method. This criterion $\varepsilon_{g l o b}$ has been written in such a way that if the solution verifies that $\varepsilon_{\text {glob }}$ is sufficiently small, then this solution provides good properties of the equation of motion and Signorini-Coulomb contact law. Consequently, this criterion remains valid for methods developed using an augmented Lagrangian (e.g., SAL and EAL methods).

This criterion can be stated as:

$$
\begin{equation*}
\varepsilon_{\mathrm{glob}}=\frac{1}{N_{c}} \sum_{c=1}^{N_{c}}\left[\varepsilon_{\text {motion }}^{c}+\varepsilon_{\text {proj }}^{c}+\varepsilon_{b_{c}}+\varepsilon_{\text {pen }}^{c}\right] \tag{37}
\end{equation*}
$$

where

- $\varepsilon_{\text {motion }}^{c}=\left\|\tilde{\mathbf{u}}^{c}-\tilde{\mathbf{u}}_{m}^{c}\right\|$ where $\tilde{\mathbf{u}}_{m}^{c}=\tilde{\mathbf{u}}^{c, i}+\sum_{\alpha=1}^{N_{c}} W_{c \alpha} \mathbf{r}^{\alpha}$, so $\varepsilon_{\text {motion }}$ measures the error on the equation of motion (see Equation (24), this term vanishes for the SBP and SAL methods);
- $\varepsilon_{\text {proj }}^{c}=\left\|\mathbf{r}^{c}-\operatorname{proj}\left(\mathbf{r}^{c}, K_{\mu}\right)\right\|$ is the error for the projection onto the Coulomb cone (vanishing for the SBP method);
- $\varepsilon_{b_{c}}=\left|\tilde{\mathbf{u}}^{c} \cdot \mathbf{r}^{c}+\mu r_{n}^{c}\left\|\tilde{\mathbf{u}}_{t}^{c}\right\|\right|$ is the absolute value of the bi-potential that has to vanish if and only if the couple ( $\tilde{\mathbf{u}}^{c}, \mathbf{r}^{c}$ ) verifies the Signorini Coulomb contact law (see formula 11);
- $\varepsilon_{p e n}^{c}=-\min \left(0, \tilde{u}_{n}^{c}\right)$ is the value of the penetration.

Remark 7
One can notice that it is absolutely necessary to verify in the criterion that there is no penetration, because nothing in the presented algorithm ensures that this condition is satisfied at the end of the loop. Moreover, if this condition is not satisfied, the other part of the bipotential can be negative or can vanish, even if the couple ( $\tilde{\mathbf{u}}, \mathbf{r}$ ) is not a solution.

## 4. NUMERICAL RESULTS

I present in this section three numerical examples with an increasing complexity.
In these computations, the descent parameter $\rho$ is taken in such a way that the result is optimal in terms of time computing. Denoting $\bar{\rho}=\frac{m_{i} m_{j}}{m_{i}+m_{j}} \frac{1}{\Delta t}$, for the SBP and the SAL methods, we have chosen $\rho=0.6 \bar{\rho}$, whereas for the EBP and the EAL methods, it is better the take $\rho=\bar{\rho}$. We recall that it has been shown that, for the bipotential method (see, e.g., [3]) and the augmented Lagrangian method (see, e.g., [11]), the parameter $\rho$ has to verify $\rho<2 \bar{\rho}$ in order to ensure the convergence. Generally, for these two methods, the convergence is very sensitive to this parameter. We will show in the last paragraph of this study that for the EBP method, the parameter $\rho$ can be taken in a large range around the value $\bar{\rho}$ without changing dramatically the convergence of the method.

At each iteration of the NLGS algorithm, the Newton algorithm is stopped either if the convergence is obtained $\left(\varepsilon_{\text {Newt }}^{c} \leqslant 10^{-5}\right)$ or if the number of iterations of the Newton algorithm reaches 100 when there is no convergence.

### 4.1. Ball sliding on a plane

In this first example, we consider a ball placed on a table with an initial horizontal velocity equal to $1.5 \mathrm{~m} \cdot \mathrm{~s}^{-1}$. The radius of the ball is equal to $5 \cdot 10^{-3} \mathrm{~m}$, and the friction coefficient between the table and the ball is equal to $\mu=0.7$. The time step of discretization is equal to $10^{-4} \mathrm{~s}$. In this experiment, the ball first slides on the table, and then, the ball rolls without sliding (Figure 4). The global stopping criterion is equal to $\varepsilon_{\text {glob }}=10^{-10}$.

We can observe from these numerical results that the error coming from the projection is very small for the four methods. The SBP method and the SAL method give very close results, both in terms of quality (Figures 5 and 6) and in terms of time computing (Table I). Nevertheless, we can notice that the time computing is smaller with the SAL method, because there is less computations at each iteration (e.g., there is no term such as $\left\|\tilde{\mathbf{u}}_{t}\right\|$ and projection is easier to compute). The EBP method provides better results, both in terms of quality (Figure 7) and in terms of time computing ( $6.5 \%$ better). The EAL method converges after the first NLGS iteration for every time step, and consequently, this is the faster method on this example ( $7 \%$ faster than the SBP method).

### 4.2. Sedimentation of four balls in a box

In this second experiment, we consider the sedimentation of four balls of radii ranging from $4 \cdot 10^{-4}$ to $5 \cdot 10^{-4} \mathrm{~m}$ (Figure 8). For the computations, the time step of discretization is equal to $\Delta t=10^{-4} \mathrm{~s}$, and the NLGS loop is stopped either if the global stopping criterion on the NLGS method is equal to $\varepsilon_{\text {glob }}=10^{-10}$ or after 5000 iterations if there is no convergence (this case never occurs in this experiment). The friction coefficient between the balls and between the balls and the walls is equal to $\mu=0.3$.


Figure 4. Example 1—A ball is launched with an initial horizontal velocity (left). First, the ball slides. Then, the ball rolls without slipping (right).


Figure 5. Example 1—Convergence for the standard bipotential-based (SBP) method, fifth iteration. The last two curves collapse. Thus, the total error is essentially due to the penetrations.


Figure 6. Example 1-Convergence for the standard augmented Lagrangian (SAL) method, fifth iteration. The last two curves collapse. Thus, the total error is essentially due to the penetrations.

Table I. Comparison of the results obtained by the four methods on the first example (after the 2000th time step).

| Method | Number of <br> NLGS iterations <br> (last time step) | Error <br> $\varepsilon_{\text {glob }}$ <br> (last time step) | Total <br> CPU time (s) |
| :--- | :---: | :---: | :---: |
| SBP | 18 | $0.384 \cdot 10^{-10}$ | 9.44 |
| SAL | 18 | $0.384 \cdot 10^{-10}$ | 9.28 |
| EBP | 1 | 0 | 8.83 |
| EAL | 1 | $0.175 \cdot 10^{-13}$ | 8.78 |

NLGS, nonlinear Gauss-Seidel; SBP, standard bipotential-based method; SAL, standard augmented Lagrangian; EBP, enhanced bipotential method; EAL, enhanced augmented Lagrangian.


Figure 7. Example 1-Convergence for the Newton and enhanced bipotential (EBP) method, fifth iteration. The curve corresponding to the penetrations does not appear because the penetrations are lower than $10^{-14}$.


Figure 8. Example 2—Sedimentation of four balls under the gravity effect.


Figure 9. Example 2-Convergence of the nonlinear Gauss-Seidel iterations for the standard bipotentialbased (SBP) method (1000th time step). The last two curves overlap, showing that the global error is governed by the error of penetration.


Figure 10. Example 2-Convergence of the nonlinear Gauss-Seidel iterations for the standard augmented Lagrangian (SAL) method (1000th time step). The last two curves overlap, showing that the global error is governed by the error of penetration.


Figure 11. Example 2-Convergence of the nonlinear Gauss-Seidel iterations for the Newton and enhanced bipotential (EBP) method (1000th time step). The last two curves overlap, which shows that the global error is governed by the error on the equations of motion.

Table II. Comparison of the results obtained by the four methods on the second example (after the 1000th time step).

| Method | Number of <br> NLGS iterations <br> (last time step) | Error <br> $\varepsilon_{\text {glob }}$ <br> (last time step) | Maximal <br> penetration <br> (last time step) | Total <br> CPU time (s) |
| :--- | :---: | :---: | :---: | :---: |
| SBP | 305 | $0.949 \cdot 10^{-12}$ | $0.310 \cdot 10^{-11}$ | 2.92 |
| SAL | 301 | $0.980 \cdot 10^{-12}$ | $0.340 \cdot 10^{-11}$ | 2.87 |
| EBP | 161 | $0.635 \cdot 10^{-12}$ | $0.641 \cdot 10^{-12}$ | 2.59 |
| EAL | 158 | $0.973 \cdot 10^{-12}$ | $0.208 \cdot 10^{-19}$ | 2.43 |

NLGS, nonlinear Gauss-Seidel; SBP, standard bipotential-based method; SAL, standard augmented Lagrangian; EBP, enhanced bipotential method; EAL, enhanced augmented Lagrangian.

Like in the previous simulations, the SBP and SAL methods provide very similar results (Figures 9 and 10). For these two methods, we can notice that here the global error is essentially due to the penetrations. The SAL method is $2 \%$ faster than the SBP method (Table II).

Results obtained by the EBP method are better (Figure 11), and here, the overall error is governed by the error on the equations of motion. The EBP method is $11 \%$ faster than the SBP method, and the penetration is five times smaller. In this example, the EAL is the faster method ( $16.8 \%$ faster than the SBP method), and the penetration is very small (Figure 12).

### 4.3. Sedimentation of 500 balls

In this example, we consider the sedimentation of 500 balls (Figure 13) of radii ranging from $2.5 \cdot 10^{-4}$ to $5 \cdot 10^{-4} \mathrm{~m}$, the time step of discretization is equal to $\Delta t=5 \cdot 10^{-5} \mathrm{~s}$, and the NLGS loop is stopped if the global estimator (37) verifies $\varepsilon_{\text {glob }} \leqslant 10^{-12}$ or after 5000 iterations if there is no convergence. The friction coefficient between the balls, and between the balls and the walls is equal to $\mu=0.3$.

The results in Table III are obtained after 1000 time steps $\left(t=5 \cdot 10^{-2} \mathrm{~s}\right)$.


Figure 12. Example 2-Convergence of the nonlinear Gauss-Seidel iterations for the Newton and enhanced augmented Lagrangian (EAL) method (1000th time step). The last two curves overlap, and the other one does not appear on the figure because the corresponding errors are lower than $10^{-16}$.


Figure 13. Example 3-Zoom-in images of balls falling under the gravity effect. Initial configuration on the left, final configuration on the right.

Table III. Comparison of the results obtained by the four methods on the third example (after the 1000th iteration, $N_{\max }=5000$ iterations).

| Method | Number of <br> NLGS iterations <br> (last time step) | Error <br> $\varepsilon_{\text {glob }}$ <br> (last time step) | Maximal <br> penetration <br> (last time step) | Total <br> CPU time (s) |
| :--- | :---: | :---: | :---: | ---: |
| SBP | 5000 | $0.119 \cdot 10^{-6}$ | $0.213 \cdot 10^{-5}$ | 1092.95 |
| SAL | 5000 | $0.135 \cdot 10^{-6}$ | $0.533 \cdot 10^{-5}$ | 973.31 |
| EBP | 5000 | $0.156 \cdot 10^{-6}$ | $0.286 \cdot 10^{-6}$ | 854.31 |
| EAL | 5000 | $0.101 \cdot 10^{-6}$ | $0.390 \cdot 10^{-17}$ | 916.65 |

NLGS, nonlinear Gauss-Seidel; SBP, standard bipotential-based method; SAL, standard augmented Lagrangian; EBP, enhanced bipotential method; EAL, enhanced augmented Lagrangian.

Table IV. Comparison of the results obtained for various values of $\rho=\alpha \bar{\rho}$ on the third example (after the 500th iteration, $N_{\max }=5000$ iterations)

| $\alpha$ | Number of NLGS <br> iterations <br> (last time step) | Maximal <br> penetration <br> (last time step) | Total CPU <br> time (s) |
| :---: | :---: | :---: | :---: |
| 5 | 652 | $0.110 \cdot 10^{-6}$ | 65.08 |
| 2 | 414 | $0.177 \cdot 10^{-6}$ | 55.86 |
| 1 | 750 | $0.149 \cdot 10^{-6}$ | 53.95 |
| $\frac{1}{2}$ | 812 | $0.634 \cdot 10^{-6}$ | 78.43 |
| $\frac{1}{5}$ | 667 | $0.219 \cdot 10^{-5}$ | 176.14 |

NLGS, nonlinear Gauss-Seidel.


Figure 14. Example 3-Convergence of the nonlinear Gauss-Seidel iterations for the standard bipotentialbased (SBP) method (1000th time step). The last two curves collapse.

In this example, the difference between methods SBP and SAL on the one hand, and the methods EBP and EAL on the other hand is larger (Table III). We can notice that the SAL method is $10.95 \%$ faster than the SBP method, and the EBP is $21.83 \%$ faster than the SBP method. Here, the EAL method is no longer the faster one, but the penetration is very small (Figure 17). Again, for the first two methods the global error is essentially due to the penetrations (Figures 14 and 15), whereas for the last two methods, the error is essentially due to the failure to follow precisely the equations of motion (Figures 16 and 17).

### 4.4. Discussion on the descent parameter $\rho$

We consider again the third example solved by the Newton and bipotential method ( $\varepsilon_{\text {tot }}=10^{-8}$, maximal number of iterations of Newton method equal to $100, \varepsilon_{\text {Newt }}=10^{-5}$, 500th time step). Here, we take $\bar{\rho}=\frac{m_{i} m_{j}}{m_{i}+m_{j}} \frac{1}{\Delta t}$, and we consider $\rho=\alpha \bar{\rho}$, for various values of $\alpha$.

These results show one of the main advantage of the EBP method. Indeed, one can notice that in Table IV, the CPU time and the quality of the solution are very similar if $\alpha$ is equal to 1 or 2 . Even if $\alpha$ is equal to 5 , the convergence is not badly damaged. One remains that for this value of $\alpha$, the SBP and the SAL methods are no longer convergent. If the parameter $\alpha$ is small, the method converges, but the convergence rate is very small. One can notice that the EAL method is much more sensitive to the parameter $\alpha$, essentially in the convergence of the Newton method. This phenomena can be explained by the fact that the EAL method differs from the SAL method only if the augmented reaction $\tau^{i+1}$ is outside the admissible set $\mathcal{C}_{c}\left(\mu \mathbf{r}^{i+1}\right)$.


Figure 15. Example 3-Convergence of the nonlinear Gauss-Seidel iterations for the standard augmented Lagrangian (SAL) method (1000th time step). The last two curves collapse.


Figure 16. Example 3-Convergence of the nonlinear Gauss-Seidel iterations for the Newton and enhanced bipotential method (EBP) (1000th time step). The last two curves collapse.


Figure 17. Example 3-Convergence of the nonlinear Gauss-Seidel iterations for the Newton and enhanced augmented Lagrangian (EAL) method (1000th time step). The last two curves collapse. The curve corresponding to the penetrations does not appear because the penetrations are lower than $10^{-14}$.

## 5. CONCLUSION

The results presented show that, using an appropriate Newton method, it is possible to improve the computational time to more than $20 \%$ compared with the standard methods . Moreover, one principal drawback of that type of methods, that is, the dependance of the results on the parameter $\rho$, does not exist anymore .

On the other hand, the stopping criterion developed appears to be efficient. It permits to have a unified criterion to compare the methods. Moreover, it ensures the mechanical quality of the solutions.

In the future, this method will be extended to the case of a contact law with adhesion. The difficulty here is the writing of appropriate local equations to model the adhesion. Even in the case of a frictional contact law, one can notice that the local equations that have to be solved come from a choice, and the efficiency of the method strongly depends on this choice.

## ACKNOWLEDGEMENTS

This work was supported by the French CNRS grant 'PEPS Opticontact' and by the project 'TRIBAL' supported by the Picardy region and the European funds FEDER.

## REFERENCES

1. De Saxcé G, Feng Z-Q. New inequality and functional for contact with friction: the implicit standard material approach. Mechanics of Structures and Machines 1991; 19:301-325.
2. Alart P, Curnier A. A mixed formulation for frictional contact problems prone to Newton like solution method. Computer Methods in Applied Mechanics and Engineering 1991; 92(3):353-375.
3. Feng Z-Q, Joli P, Cros J-M, Magnain B. The bi-potential method applied to the modeling of dynamics problems with friction. Computational Mechanics 2005; 36(5):375-383.
4. Klarbring A. Mathematical programming in contact problems. In Computational Methods in Contact Mechanics, Aliabdali MH, Brebbia CA (eds). Computational Mechanics. Publications: Southampton, 1993; 233-263.
5. Wriggers P. Finite element algorithms for contact problems. Archives of Computational Methods in Engineering 1995; 2:1-49.
6. Jean M, Moreau JJ. Unilaterality and dry friction in the dynamics of rigid bodies collection. In Contact Mechanics International Symposium, Curnier A (ed.). Presses Polytechniques et Universitaires Romanes: Lausanne, 1992; 31-48.
7. Jean M. The non smooth contact dynamics method. Computer Methods Applied Mathematics Engineering 1999; 177:235-257.
8. Moreau JJ. Unilateral contact and dry friction in finite freedom dynamics. In Non Smooth Mechanics and Applications, CISM Courses and Lectures, Vol. 302, Moreau JJ, Panagiotopoulos P-D (eds). Springer-Verlag, Wien: New York, 1988; 1-82.
9. Moreau JJ. Some numerical methods in multibody dynamics: application to granular materials. European Journal of Mechanics - A/Solids 1994; 13(4-suppl.):93-114.
10. Moreau JJ. Numerical aspect of sweeping process. Computer Methods Applied Mathematics Engineering 1999; 177:329-349.
11. Renouf M, Alart P. Conjugate gradient type algorithms for frictional multicontact problems: applications to granular materials. Computer Methods in Applied Mechanics and Engineering 2004; 194(18-20):2019-2041.
12. Jourdan F, Alart P, Jean M. A Gauss-Seidel like algorithm to solve frictional contact problems. Computer Methods in Applied Mechanics an Engineering 1998; 155:31-47.
13. Renouf M, Dubois F, Alart P. A parallel version of the non smooth contact dynamics algorithm applied to the simulation of granular media. Journal of Computational and Applied Mathematics 2004; 168:375-38.
14. Fortin J. Simulation numérique de la dynamique des systèmes multicorps appliquée aux milieux granulaires. PhD Thesis (in french), University of Lille 1, 2000.
15. Fortin J, Millet O, de Saxcé G. Numerical simulation of granular materials by an improved discrete element method. International Journal for Numerical Methods in Engineering 2005; 62(5):639-663.
16. Sanni I. Modélisation et simulation bi et tri-dimensionnelles de la dynamique unilatérale des systèmes multi-corps de grandes tailles: application aux milieux granulaires. PhD Thesis (in french), University of Picardy, 2006.
17. Joli P, Feng Z-Q. Uzawa and Newton algorithms to solve frictional contact problems within the bi-potential framework. International Journal for Numerical Methods in Engineering 2008; 73(3):317-330.
18. Ramaniraka NA. Thermomécanique des contacts entre deux solides déformables. PhD thesis no. 1651 (in french), Ecole Polytechnique Fédérale de Lausanne, 1991.
19. Renouf M, Alart P. Gradient type algorithms for $2 \mathrm{~d} / 3 \mathrm{~d}$ frictionless/frictional multicontact problems. In ECCOMAS 2004, Majava K, Neittaanmaki P, Rossi T, Pironneau O, Owen R, Mikkola M (eds): Jyvaskyla, July 2004; 24-28.
20. Renouf M, Acary V, Dumont G. Comparison of algorithms for collisions, contact and friction in view of real-time applications. In Multibody Dynamics 2005 proceedings, International Conference on Advances in Computational Multibody, Goicolea JM, Cuadrado J, García Orden JC (eds), Madrid, June 2005; 21-24.
21. Renouf M, Bonamy D, Dubois F, Alart P. Numerical simulation of two-dimensional steady granular flows in rotating drum: on surface flow rheology. Physics of Fluids 2005; 17:103-303.
22. Rockafellar RT. Convex Analysis, Vol. 28, Princeton Math. Series. Princeton Univ. Press: Princeton, 1970.
23. Fortin J, Hjiaj M, de Saxcé G. An improved discrete element method based on a variational formulation of the contact law. Computers and Geotechnics 2002; 29:609-640.

[^0]:    *Correspondence to: S. Dumont, LAMFA, Université de Picardie Jules Verne - CNRS UMR 7352, 33, rue Saint-Leu, 80000 Amiens, France.
    ${ }^{\dagger}$ E-mail: serge.dumont@u-picardie.fr

