Convergence order of a numerical scheme for sweeping process

Frédéric Bernicot
CNRS - Université Lille 1
59655 Villeneuve d'Ascq Cedex
frederic.bernicot@math.univ-lille1.fr

Juliette Venel
LAMAV, Université de Valenciennes et du Hainaut-Cambrésis
Mont Houy 59313 Valenciennes Cedex 9
juliette.venel@univ-valenciennes.fr

September 16, 2010

Contents

Abstract

In this paper, we show that the convergence order of the numerical scheme introduced in [12] for sweeping processes is equal to one. The considered differential inclusions involve a set-valued map, given by a finite number of constraints. The proof rests on a metric qualification condition between the sets associated with each constraint.

Key-words: Differential inclusions - Subdifferential calculus - Numerical analysis - Prediction-correction algorithm.

MSC: 34A60, 65L20.

1 Introduction

In [12], an implementable scheme was introduced to compute the discretized solutions of some sweeping processes. Its convergence was proved by using compactness arguments. The aim of this paper is to specify the convergence order of this scheme by directly estimating the approximation error.

Let us briefly recall the mathematical framework. A problem of perturbated sweeping process is a first order differential inclusion which can be expressed as follows

$$\frac{d\mathrm{q}}{dt}(t) + \mathrm{N}(Q(t), \mathrm{q}(t)) \ni \mathrm{f}(t, \mathrm{q}(t)), \tag{1}$$

where $Q(\cdot)$ is a set-valued map, N denotes the proximal normal cone (see Definition 2.1) and f represents a perturbation. We refer the reader to [9, 2, 3] for an overview of this topic.

The current work falls within the realm introduced in [12] (recalled in Section 2). We deal with a set-valued map $Q(\cdot)$, defined by a finite number of smooth convex constraints (more precisely for all t, the set $Q(t)$ is the intersection of complements of convex sets). The considered numerical scheme is based on a local approximation of $Q(t)$ by convex sets, which makes this algorithm implementable.

In this context, we prove the following result:

Theorem 1.1. *There exists a constant* $C_0 > 0$ *such that for* h *small enough*

$$\|\mathrm{q}_h - \mathrm{q}\|_{L^\infty([0,T])} \le C_0 h,$$

where q *and* q_h *are the continuous and discrete solutions of (1).*

We emphasize that this new approach allows us to go around the compactness arguments, used in [12]. So it permits to extend the convergence result of[12] in an infinite dimensional Hilbert space (see Remark 3.5).

The paper is structured as follows: In Section 2, we describe the mathematical framework by specifying notations and assumptions which will be used throughout the paper. Then in Section 3, after recalling the prediction-correction scheme proposed in [12], we prove in Theorem 3.4 that the discretized solution converges to the exact solution with order 1 . This proof rests on a metric qualification condition which is checked in Section 4.

2 Preliminaries

In the sequel, the space $\mathbb{R}^d$ is equipped with its Hilbertian structure. We write $B(x, r)$ for the closed ball of center $x \in \mathbb{R}^d$ and radius $r > 0$.
We consider perturbed sweeping process by a set-valued map $Q : [0, T] \rightrightarrows \mathbb{R}^d$ satisfying that for every $t \in [0, T]$, $Q(t)$ is the intersection of complements of smooth convex sets. Let us first specify the set-valued map Q. For $i \in \{1, \dots, p\}$, let $g_i : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ be a convex function with respect to the second variable. For every $t \in [0, T]$, we introduce the sets $Q_i(t)$ defined by

$$Q_i(t) := \left\{ \mathrm{q} \in \mathbb{R}^d\, ,\ g_i(t, \mathrm{q}) \ge 0 \right\}, \tag{2}$$

and the feasible set $Q(t)$ (supposed to be nonempty) is

$$Q(t) := \bigcap_{i=1}^{p} Q_i(t). \tag{3}$$

The associated perturbed sweeping process can be expressed as follows:

$$\begin{cases} \dfrac{d\mathrm{q}}{dt}(t) + \mathrm{N}(Q(t), \mathrm{q}(t)) \ni \mathrm{f}(t, \mathrm{q}(t)) \text{ for a.e. } t \in [0, T] \\ \mathrm{q}(0) = \mathrm{q}_0 \in Q(0). \end{cases} \tag{4}$$

We write $\mathrm{N}(Q(t), q(t))$ for the proximal normal cone to $Q(t)$ at $q(t)$, below defined.

Definition 2.1 ([1]). *Let S be a closed subset of $\mathbb{R}^d$.*
The proximal normal cone to S at $\mathbf{x}$ is defined by:

$$\mathrm{N}(S,\mathbf{x}) := \left\{\mathbf{v} \in \mathbb{R}^d,\ \exists\alpha > 0,\ \mathbf{x} \in \mathrm{P}_S(\mathbf{x}+\alpha\mathbf{v})\right\},$$

where

$$\mathrm{P}_S(\mathbf{y}) := \{\mathbf{z} \in S,\ d_S(\mathbf{y}) = |\mathbf{y}-\mathbf{z}|\}, \qquad \textit{with} \quad d_S(\mathbf{y}) := \inf_{\mathbf{z}\in S} |\mathbf{y}-\mathbf{z}|$$

corresponds to the Euclidean projection onto S.

This differential inclusion can be thought as follows: the point $\mathrm{q}(t)$, submitted to the perturbation $\mathrm{f}(t,\mathrm{q}(t))$, has to stay in the feasible set $Q(t)$. To obtain well-posedness results for (4), we will make the following assumptions which ensure the uniform prox-regularity of $Q(t)$ for all $t \in [0,T]$. We suppose there exists $c > 0$ and for all t in $[0,T]$ open sets $U_i(t) \supset Q_i(t)$ such that

$$d_H(Q_i(t), \mathbb{R}^d \setminus U_i(t)) > c, \tag{A0}$$

where d_H denotes the Hausdorff distance. We set $U(t) := \bigcap_{i=1}^p U_i(t)$. Moreover we assume there exist constants $\alpha, \beta, M > 0$ such that for all t in $[0,T]$, $g_i(t,\cdot)$ belongs to $C^2(U_i(t))$ and satisfies

$$\forall\, \mathrm{q} \in U_i(t)\,,\ \alpha \le |\nabla_{\mathrm{q}} g_i(t,\mathrm{q})| \le \beta, \tag{A1}$$

$$\forall\, \mathrm{q} \in \mathbb{R}^d\,,\ |\partial_t g_i(t,\mathrm{q})| \le \beta, \tag{A2}$$

$$\forall\, \mathrm{q} \in U_i(t)\,,\ |\partial_t \nabla_{\mathrm{q}} g_i(t,\mathrm{q})| \le M, \tag{A3}$$

and

$$\forall\, \mathrm{q} \in U_i(t)\,,\ |\mathrm{D}^2_{\mathrm{q}} g_i(t,\mathrm{q})| \le M. \tag{A4}$$

For all $t \in [0,T]$ and $\mathrm{q} \in Q(t)$, we denote by $I(t,\mathrm{q})$ the active set at q

$$I(t,\mathrm{q}) := \{i \in \{1,\ldots,p\}\,,\ g_i(t,\mathrm{q}) = 0\} \tag{5}$$

and for every $\rho > 0$, we put:

$$I_\rho(t,\mathrm{q}) := \{i \in \{1,\ldots,p\}\,,\ g_i(t,\mathrm{q}) \le \rho\}\,. \tag{6}$$

In addition we assume there exist $\gamma > 0$ and $\rho > 0$ such that for all $t \in [0,T]$,

$$\forall\, \mathrm{q} \in Q(t)\,,\ \forall\, \lambda_i \ge 0,\ \sum_{i\in I_\rho(t,\mathrm{q})} \lambda_i |\nabla_{\mathrm{q}}\, g_i(t,\mathrm{q})| \le \gamma \left| \sum_{i\in I_\rho(t,\mathrm{q})} \lambda_i \nabla_{\mathrm{q}}\, g_i(t,\mathrm{q}) \right|. \tag{A5}$$

We will use the following weaker assumption:

$$\forall\, \mathrm{q} \in Q(t)\,,\ \forall\, \lambda_i \ge 0,\ \sum_{i\in I(t,\mathrm{q})} \lambda_i |\nabla_{\mathrm{q}}\, g_i(t,\mathrm{q})| \le \gamma \left| \sum_{i\in I(t,\mathrm{q})} \lambda_i \nabla_{\mathrm{q}}\, g_i(t,\mathrm{q}) \right|. \tag{A5'}$$

In particular, this last assumption implies that for all t, the gradients of the active inequality constraints $\nabla_{\mathrm{q}}\, g_i(t,\mathrm{q})$ are positive-linearly independent at all $\mathrm{q} \in Q(t)$, which is usually called the Mangasarian-Fromowitz constraint qualification (MFCQ). Conversely the MFCQ condition at a point q yields a local version of Inequality (A5').

We recall some useful results established in [12] (Propositions 2.8, 2.9, 2.11 and Theorem 2.12 in [12]).

Proposition 2.2. *For all $t \in [0,T]$ and $\mathrm{q} \in Q(t)$,*

$$\mathrm{N}(Q(t),\mathrm{q}) = \sum \mathrm{N}(Q_i(t),\mathrm{q}) = -\sum_{i\in I(t,\mathrm{q})} \mathbb{R}^+ \nabla_{\mathrm{q}}\, g_i(t,\mathrm{q}).$$

Proposition 2.3. *Under assumptions (A1), (A4) and (A5'), the set $Q(t)$ is η-prox-regular with $\eta = \dfrac{\alpha}{M\gamma}$, for every $t \in [0,T]$.*

Proposition 2.4. *Under assumptions (A0), (A1), (A2) and (A5), the set-valued map Q is Lipschitz continuous with respect to the Hausdorff distance. More precisely there exists $K_L > 0$ such that*

$$\forall t,s \in [0,T]\,,\ d_H(Q(t),Q(s)) \le K_L|t-s|.$$

Theorem 2.5. *Let $T > 0$ and $\mathrm{f} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be a measurable map satisfying:*

$$\exists K_{\mathrm{f}} > 0\,,\ \forall \mathrm{q} \in \bigcup_{s\in[0,T]} Q(s)\,,\ \forall t \in [0,T]\,,\ |\mathrm{f}(t,\mathrm{q}) - \mathrm{f}(t,\tilde{\mathrm{q}})| \le K_{\mathrm{f}}|\mathrm{q}-\tilde{\mathrm{q}}| \tag{7}$$

$$\exists L_{\mathrm{f}} > 0\,,\ \forall \mathrm{q} \in \bigcup_{s\in[0,T]} Q(s)\,,\ \forall t \in [0,T]\,,\ |\mathrm{f}(t,\mathrm{q})| \le L_{\mathrm{f}}(1+|\mathrm{q}|). \tag{8}$$

Then, under Assumptions (A0), (A1), (A2), (A4) and(A5) for all $\mathrm{q}_0 \in Q(0)$, the following problem

$$\begin{cases} \dfrac{d\mathrm{q}}{dt}(t) + \mathrm{N}(Q(t),\mathrm{q}(t)) \ni \mathrm{f}(t,\mathrm{q}(t)) \text{ for a.e. } t \in [0,T] \\ \mathrm{q}(0) = \mathrm{q}_0, \end{cases} \tag{9}$$

has one and only one absolutely continuous solution q *satisfying* $\mathrm{q}(t) \in Q(t)$ *for every* $t \in [0,T]$.

3 Time-stepping scheme

Let us detail the numerical scheme proposed in [12] to approximate the solution of (9) on the time interval $[0,T]$. Let $n \in \mathbb{N}^\star$, $h = T/n$ be the time step and $t_k^n = kh$ be the computational times. We denote by q_k^n the approximation of $\mathrm{q}(t_k^n)$ with $\mathrm{q}_0^n = \mathrm{q}_0$. The next configuration is computed as follows:

$$\mathrm{q}_{k+1}^n := \mathrm{P}_{\tilde{Q}(t_{k+1}^n,\mathrm{q}_k^n)}(\mathrm{q}_k^n + h\mathrm{f}(t_k^n,\mathrm{q}_k^n)) \tag{10}$$

with

$$\tilde{Q}(t,\mathrm{q}) := \{\tilde{\mathrm{q}} \in \mathbb{R}^d\,,\ g_i(t,\mathrm{q}) + \langle \nabla_{\mathrm{q}}\, g_i(t,\mathrm{q}), \tilde{\mathrm{q}} - \mathrm{q}\rangle \ge 0 \quad \forall i\} \text{ for } \mathrm{q} \in U(t) := \bigcap_{i=1}^{p} U_i(t).$$

We recall that all the gradients $\nabla_{\mathrm{q}}\, g_i(t,\mathrm{q})$ are well-defined provided that $\mathrm{q} \in U(t)$. Indeed it can be checked that this scheme is well-defined (more precisely $\tilde{Q}(t_{k+1}^n,\mathrm{q}_k^n) \subset Q(t_{k+1}^n) \subset U(t_{k+2}^n)$) for $h < \frac{c}{K_L}$ with c and K_L respectively given by Assumption (A0) and Proposition 2.4 (see Proposition 3.1 in [12]). Thus every computed configuration is feasible and the set $\tilde{Q}(t,\mathrm{q})$ can be seen as an inner convex approximation of $Q(t)$ with respect to q.
This scheme is a prediction-correction algorithm: predicted position vector $\mathrm{q}_k^n + h\mathrm{f}(t_k^n,\mathrm{q}_k^n)$, that may not be admissible, is projected onto the approximate set of feasible configurations.

Before stating the result of convergence, we introduce some notations. We define the piecewise constant function f^n as follows,

$$\mathrm{f}^n(t) = \mathrm{f}(t_k^n,\mathrm{q}_k^n) \;\text{ if }\; t \in [t_k^n, t_{k+1}^n[,\ k < n \;\text{ and }\; \mathrm{f}^n(T) = \mathrm{f}(t_{n-1}^n,\mathrm{q}_{n-1}^n). \tag{11}$$

We denote by q^n the continuous, piecewise linear function satisfying for $k \in \{0, \dots, n\}$, $\mathrm{q}^n(t_k^n) = \mathrm{q}_k^n$. To finish, we introduce the functions ρ and θ defined by

$$\rho^n(t) = t_k^n \text{ and } \theta^n(t) = t_{k+1}^n \text{ if } t \in [t_k^n, t_{k+1}^n[,\ \rho^n(T) = T \text{ and } \theta^n(T) = T.$$

We recall some results about these approximate solutions (see Subsection 3.2 in [12] for details) :

Theorem 3.1. *Let suppose that for all* $\mathrm{q} \in \bigcup_{s\in[0,T]} Q(s)$,

$$\mathrm{f}(\cdot, \mathrm{q}) \quad \textit{is Riemann-integrable on } [0, T]. \tag{12}$$

Then with the assumptions of Theorem 2.5, q^n *tends to* q *in* $C^0([0, T], \mathbb{R}^d)$*, where* $t \mapsto \mathrm{q}(t)$ *is the unique solution of (9).*

Remark 3.2. *If we replace the definition (11) of* f^n *with*

$$\mathrm{f}^n(t) = \frac{1}{h}\int_{t_k^n}^{t_{k+1}^n} \mathrm{f}(s, \mathrm{q}_k^n)ds \quad \textit{if} \quad t \in [t_k^n, t_{k+1}^n[,\ k < n \quad \textit{and} \quad \mathrm{f}^n(T) = \frac{1}{h}\int_{T-h}^{T} \mathrm{f}(s, \mathrm{q}_{n-1}^n)ds,$$

the hypothesis (12) is unnecessary.

Proposition 3.3. *There exists* $C, D, K > 0$ *such that*

$$\sup_n \|\mathrm{q}^n\|_{L^\infty([0,T])} \leq C, \quad \sup_n \left\| \frac{d\mathrm{q}^n}{dt} \right\|_{L^\infty([0,T])} \leq K$$

and for n *large enough,*

$$d_{\tilde{Q}(t_{k+1}^n, \mathrm{q}_k^n)}(\mathrm{q}_k^n) \leq Dh. \tag{13}$$

We now come to the main result of the present paper which specifies the convergence order of the previous scheme.

Theorem 3.4. *There exists a constant* $C_0 > 0$ *such that for* n *large enough*

$$\|\mathrm{q}^n - \mathrm{q}\|_{L^\infty([0,T])} \leq C_0 \frac{T}{n},$$

where q *is the solution of (4).*

Proof. We check that the sequence $(\mathrm{q}^n)_n$ is of Cauchy type.
Let $m \geq n$ be large enough. Since for $k \in \{0, .., n-1\}$

$$\mathrm{q}_{k+1}^n = \mathrm{P}_{\tilde{Q}(t_{k+1}^n, \mathrm{q}_k^n)}(\mathrm{q}_k^n + h\mathrm{f}(t_k^n, \mathrm{q}_k^n))$$

and $\tilde{Q}(t_{k+1}^n, \mathrm{q}_k^n)$ is a closed convex set, it comes : for all $y \in \mathbb{R}^d$

$$\langle \mathrm{q}_k^n + h\mathrm{f}(t_k^n, \mathrm{q}_k^n) - \mathrm{q}_{k+1}^n, y - \mathrm{q}_{k+1}^n \rangle \leq |\mathrm{q}_k^n + h\mathrm{f}(t_k^n, \mathrm{q}_k^n) - \mathrm{q}_{k+1}^n| d_{\tilde{Q}(t_{k+1}^n, \mathrm{q}_k^n)}(y). \tag{14}$$

By Assumption (8) and Proposition 3.3, we get for almost $t \in [0, T]$,

$$\left| \mathrm{f}^n(t) - \frac{d\mathrm{q}^n}{dt}(t) \right| \leq L_{\mathrm{f}}(1 + C) + K. \tag{15}$$

Consequently by dividing (14) by h, we obtain for all $y \in \mathbb{R}^d$

$$-\left\langle \frac{d\mathrm{q}^n}{dt}(t) - \mathrm{f}^n(t), y - \mathrm{q}^n(\theta^n(t)) \right\rangle \leq C_1 d_{\tilde{Q}(\theta^n(t), \mathrm{q}^n(\rho^n(t)))}(y)$$

with $C_1 := L(1+C)+K$.
Taking $y = \mathrm{q}^m(\theta^m(t))$, it follows

$$-\left\langle \frac{d\mathrm{q}^n}{dt}(t) - \mathrm{f}^n(t), \mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t)) \right\rangle \leq C_1 d_{\tilde{Q}(\theta^n(t),\mathrm{q}^n(\rho^n(t))}(\mathrm{q}^m(\theta^m(t))).$$

First case : $|\mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\rho^n(t))| \leq r/8$ (with r introduced in Theorem 4.9).
Let us chose $w \in \mathrm{P}_{Q(\theta^n(t))}(\mathrm{q}^m(\theta^m(t)))$. Hence $w \in Q(\theta^n(t))$ and

$$|w - \mathrm{q}^n(\rho^n(t))| \leq |w - \mathrm{q}^m(\theta^m(t))| + \frac{r}{8} \leq d_H(Q(\theta^n(t)), Q(\theta^m(t))) + \frac{r}{8} \leq \frac{K_L}{n} + \frac{r}{8} \leq \frac{r}{4}$$

by Proposition 2.4, for n large enough. Moreover $\mathrm{q}^n(\rho^n(t)) \in Q(\rho^n(t)) \subset U(\theta^n(t))$ for $n > K_L/c$ and Inequality (13) implies that

$$d_{\tilde{Q}(\theta^n(t),\mathrm{q}^n(\rho^n(t))}(\mathrm{q}^n(\rho^n(t))) \leq \frac{r}{4}$$

for n large enough. Then, by Theorem 4.9 and Proposition 3.6, we deduce

$$d_{\tilde{Q}(\theta^n(t),\mathrm{q}^n(\rho^n(t))}(w) \leq \kappa |\mathrm{q}^n(\rho^n(t)) - w|^2$$

where $\kappa := \Theta p M/(2\alpha)$. Hence with Propositions 3.3 and 2.4

$$\begin{aligned} d_{\tilde{Q}(\theta^n(t),\mathrm{q}^n(\rho^n(t))}(\mathrm{q}^m(\theta^m(t))) &\leq \kappa |\mathrm{q}^n(\rho^n(t)) - w|^2 + |\mathrm{q}^m(\theta^m(t)) - w| \\ &\leq \kappa |\mathrm{q}^n(\rho^n(t)) - w|^2 + d_H(Q(\theta^n(t)), Q(\theta^m(t))) \\ &\leq 2\kappa |\mathrm{q}^n(t) - \mathrm{q}^m(t)|^2 + \frac{K_L}{n} + 2\kappa \left(\frac{K+K_L}{n} \right)^2. \end{aligned}$$

Finally,

$$-\left\langle \frac{d\mathrm{q}^n}{dt}(t) - \mathrm{f}^n(t), \mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t)) \right\rangle \leq 2C_1\kappa |\mathrm{q}^n(t) - \mathrm{q}^m(t)|^2 + \frac{C_2}{n},$$

with $C_2 := C_1(K_L + 2\kappa(K+K_L)^2)$.

Second case : $|\mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\rho^n(t))| \geq r/2$.
Then by (15),

$$\begin{aligned} -\left\langle \frac{d\mathrm{q}^n}{dt}(t) - \mathrm{f}^n(t), \mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t)) \right\rangle &\leq C_1 |\mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t))| \\ &\leq \frac{2}{r} C_1 |\mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t))|^2 \\ &\leq \frac{4}{r} C_1 |\mathrm{q}^m(t) - \mathrm{q}^n(t)|^2 + \frac{4}{r} C_1 \left(\frac{2K}{n} \right)^2. \end{aligned}$$

End of the proof :
By setting $C_3 := \max\{2C_1\kappa, 4C_1/r\}$ and $C_4 := \max\{C_2, \frac{16 C_1 K^2}{r}\}$, we get

$$-\left\langle \frac{d\mathrm{q}^n}{dt}(t) - \mathrm{f}^n(t), \mathrm{q}^m(\theta^m(t)) - \mathrm{q}^n(\theta^n(t)) \right\rangle \leq C_3 |\mathrm{q}^m(t) - \mathrm{q}^n(t)|^2 + \frac{C_4}{n}.$$

By summing the previous inequality and the other one obtained by changing the role of n and m, it yields

$$\left\langle \frac{d\mathrm{q}^m}{dt}(t) - \frac{d\mathrm{q}^n}{dt}(t), \mathrm{q}^m(t) - \mathrm{q}^n(t) \right\rangle \leq (2C_3 + K_\mathrm{f})|\mathrm{q}^m(t) - \mathrm{q}^n(t)|^2 + \frac{C_5}{n}$$

where L_f is introduced in (7) and $C_5 := 2C_4 + 4K^2(K_\mathrm{f} + 1)$.
By applying Gronwall's Lemma, we have

$$\|\mathrm{q}^m - \mathrm{q}^n\|_{L^\infty}([0,T]) \leq \frac{C_4}{n} \exp((2C_3 + L_\mathrm{f})T).$$

Then, we conclude the proof by taking the limit for $m \to \infty$. □

Remark 3.5. *This proof allows us to get around the compactness arguments employed in [12] to obtain the convergence of* q_h. *Consequently, this result can be extended to the Hilbertian case. Then it can be checked that the limit satisfies the differential inclusion (4) by following the same reasoning as in [12]. In particular, we also find the existence of solutions again. However the uniqueness requires the uniform prox-regularity of sets* $Q_i(t)$ *(which implies the uniform prox-regularity of sets* $Q(t)$*).*

It remains to prove Proposition 3.6 and Theorem 4.9. We now check the first result whereas the second one will be established in the next Section.

Proposition 3.6. *For all* $t \in [0,T]$, $\mathrm{q}_0 \in U(t)$ *and all* $\mathrm{q} \in Q(t)$, *we have for all* $i \in \{1, ..., p\}$

$$d_{\tilde{Q}_i(t,\mathrm{q}_0)}(\mathrm{q}) \leq \frac{M}{2\alpha}|\mathrm{q} - \mathrm{q}_0|^2. \tag{16}$$

Proof. Let consider $i \in \{1, ..., p\}$, $\mathrm{q}_0 \in U(t)$ and $\mathrm{q} \in Q(t) \subset Q_i(t)$. We assume that $\mathrm{q} \notin \tilde{Q}_i(t, \mathrm{q}_0)$ (otherwise (16) obviously holds).
For $\ell \geq 0$, we define

$$z(\ell) := \mathrm{q} + \ell \nabla g_i(t, \mathrm{q}_0).$$

The point $z(\ell)$ belongs to $\tilde{Q}_i(t, \mathrm{q}_0)$ if and only if

$$g_i(t, \mathrm{q}_0) + \langle \nabla g_i(t, \mathrm{q}_0), \mathrm{q} - \mathrm{q}_0 \rangle + \ell |\nabla g_i(t, \mathrm{q}_0)|^2 \geq 0,$$

which is equivalent to

$$\ell \geq \ell_0 := -\frac{g_i(t, \mathrm{q}_0) + \langle \nabla g_i(t, \mathrm{q}_0), \mathrm{q} - \mathrm{q}_0 \rangle}{|\nabla g_i(t, \mathrm{q}_0)|^2} \geq 0.$$

Thus,

$$\begin{aligned} d_{\tilde{Q}_i(t,\mathrm{q}_0)}(\mathrm{q}) &\leq |\mathrm{q} - z(\ell_0)| \leq \ell_0 |\nabla g_i(t, \mathrm{q}_0)| \\ &\leq -\frac{g_i(t, \mathrm{q}_0) + \langle \nabla g_i(t, \mathrm{q}_0), \mathrm{q} - \mathrm{q}_0 \rangle}{|\nabla g_i(t, \mathrm{q}_0)|} \\ &\leq \frac{1}{|\nabla g_i(t, \mathrm{q}_0)|} \int_0^1 s \mathrm{D}^2_\mathrm{q} g_i(t, \mathrm{q}_0 + s(\mathrm{q} - \mathrm{q}_0))(\mathrm{q} - \mathrm{q}_0, \mathrm{q} - \mathrm{q}_0) ds, \end{aligned}$$

because $g_i(t, \mathrm{q}) \geq 0$. We conclude to (16) by Assumptions (A1) and (A4). □

4 Metric qualification condition

This section is devoted to the proof of Theorem 4.9, which corresponds to a metric qualification condition for the sets $\tilde{Q}_i$. Aiming that, we recall some notions of subdifferential calculus.

Definition 4.1 (proximal subdifferential)**.** *Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ be a lower semicontinuous function which is finite at $x \in \mathbb{R}^d$. The proximal subdifferential of f at x is defined by:*

$$\partial^P f(x) := \left\{ x^\star \in \mathbb{R}^d,\ \exists \alpha, \beta > 0,\ \forall |h| \leq \beta,\ f(x+h) - f(x) \geq \langle x^\star, h \rangle - \alpha |h|^2 \right\}.$$

Definition 4.2 (limiting subdifferential)**.** *Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ be a lower semicontinuous function which is finite at $x \in \mathbb{R}^d$. The limiting (or Mordukhovich) subdifferential of f at x is defined by*

$$\partial^L f(x) := \left\{ x^\star \in \mathbb{R}^d,\ x^\star = \lim_{k \to \infty} x_k^\star \quad \text{with} \quad x_k^\star \in \partial^P f(x_k),\ x_k \to x \quad \text{and} \quad f(x_k) \to f(x) \right\}.$$

Definition 4.3 (Clarke subdifferential)**.** *Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ be a Lipschitzian function. The Clarke subdifferential $\partial^C f(x)$ of f at x can be defined (see [1]) as the closed convex hull of the limiting subdifferential :*

$$\partial^C f(x) := \overline{\mathit{conv}\ \partial^L f(x)}.$$

This notion has been extended for less regular functions, we refer the reader to [11] for details.

The following property is a special case of the exact sum rule for the Clarke subdifferential (see Theorem 2 of [11]):

Lemma 4.4 (Optimality property)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function and $\phi : \mathbb{R}^d \to \mathbb{R}$ a convex Lipschitz function. If $x \in \mathbb{R}^d$ is a finite local minimum of $f + \phi$ then*

$$0 \in \partial^C f(x) + \partial^C \phi(x).$$

Let us recall the variational principle of Ekeland (see [4]).

Proposition 4.5 (Ekeland variational principle)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous which is bounded from below. Let $\epsilon > 0$ and $x \in \mathbb{R}^d$ such that*

$$\inf f \leq f(x) \leq \inf f + \epsilon.$$

Then for all $\lambda > 0$, there exists $w \in \mathbb{R}^d$ satisfying

- $f(w) \leq f(x)$
- $|x - w| \leq \lambda$
- *for all $z \neq w$, $f(z) > f(w) - \frac{\epsilon}{\lambda} |z - w|$.*

The following result comes from Theorem 2.1 in [5]. For an easy reference, we detail the proof.

Lemma 4.6. *Let $f : \mathbb{R}^d \to \mathbb{R}^+ \cup \{+\infty\}$ be a lower semi-continuous function and x_0 with $f(x_0) = 0$. Assume there exist $\gamma, \delta > 0$ such that for all*

$$x^\star \in \bigcup_{\substack{x \in B(x_0, 2\delta) \\ f(x) > 0}} \partial^C f(x)$$

we have $|x^\star| \geq \gamma$. Then for all $x \in B(x_0, \delta)$, $d_{\{f=0\}}(x) \leq \gamma^{-1} f(x)$.

Proof. Let $x \in B(x_0, \delta)$. If $f(x) \geq \gamma\delta$, then

$$d_{\{f=0\}}(x) \leq |x - x_0| \leq \delta \leq \gamma^{-1} f(x).$$

Now, we assume that $0 < f(x) < \gamma\delta$ and we set $\epsilon := f(x)$. Applying the variational principle of Ekeland (see Proposition 4.5) to f with ϵ and any $\lambda \in]\gamma^{-1}\epsilon, \delta[$. There exists $w = w(\lambda) \in \mathbb{R}^d$ such that $f(w) \leq f(x)$, $|x - w| \leq \lambda$ and

$$\forall z \neq w, \quad f(z) > f(w) - \frac{\epsilon}{\lambda}|w - z|.$$

Consequently, w minimizes $f + \epsilon\lambda^{-1}|\cdot - w|$ and by Lemma 4.4 it comes

$$0 \in \partial^C f(w) + \partial^C \psi(w)$$

where $\psi(\cdot) = \epsilon\lambda^{-1}|\cdot - w|$.
So there exists $x^\star \in \partial^C f(w)$ with $|x^\star| \leq \epsilon\lambda^{-1} < \gamma$. That is in contradiction with the assumptions as $|w - x_0| \leq |w - x| + |x - x_0| \leq 2\delta$ and so we deduce that necessarily $f(w) = 0$. Then we conclude to the desired result, since

$$d_{\{f=0\}}(x) \leq |x - w| \leq \lambda$$

holds for every $\lambda \in]\gamma^{-1}\epsilon, \delta[$. □

Frow now on, we come back to the framework of the previous sections and prove the metric qualification condition of sets $\tilde{Q}_i$.

In the sequel, we introduce convex sets C_i for $i \in \{1, ..., p\}$ and their intersection $C = \bigcap_{i=1}^{p} C_i$.
We consider the following set-valued map F

$$F : \begin{cases} \mathbb{R}^d & \rightrightarrows & \mathbb{R}^{dp} \\ x & \mapsto & F(x) := (C_1 - x) \times \cdots \times (C_p - x). \end{cases} \tag{17}$$

Let us note that $0 \in F(x)$ if and only if $x \in C$.

Proposition 4.7. *Consider the function f defined by $f(x) := d_{F(x)}(0)$ where F is given by (17). The map f is Lipschitzian and for all $x \notin C$,*

$$\partial^P f(x) \subset \partial^C f(x) = \left\{ \sum_{i,\ x \notin C_i} \frac{y_i}{|y|} \right\},$$

where $y = \mathrm{P}_{F(x)}(0)$. In other words, for all $i \in \{1, ..., p\}$, $y_i + x \in \mathrm{P}_{C_i}(x)$, hence $-y_i \in \mathrm{N}(C_i, x + y_i)$.

Proof. For all $x \in \mathbb{R}^d$,

$$f(x) = d_{F(x)}(0) = d_\Pi(\phi(x))$$

where $\Pi := \otimes_{i=1}^{p} C_i$ and $\phi(x) := (x, \ldots, x) \in \mathbb{R}^{dp}$. For $x \notin C$,

$$\partial^C f(x) = \partial^C (d_\Pi \circ \phi)(x) = {}^t(1, \ldots, 1) \cdot \partial^C d_\Pi(\phi(x))$$

thanks to Corollary 1 in [11]. By convexity of the sets C_i, d_Π is a convex functions and so

$$\partial^C f(x) = {}^t(1, \ldots, 1) \cdot \partial^P d_\Pi(\phi(x)),$$

see Remark 4.8. First we claim that

$$\partial^P d_\Pi(\phi(x)) \subset [\otimes_{i=1}^p \mathcal{E}_i(x)] \bigcap S(0,1), \tag{18}$$

with $\mathcal{E}_i(x) := \frac{d_{C_i}(x)}{d_\Pi(\phi(x))}\partial^P d_{C_i}(x)$ if $x \notin C_i$ and $\mathcal{E}_i(x) := \{0\}$ else.
Indeed, let $x^\star$ belong to $\partial^P d_\Pi(\phi(x))$. By definition, for some $\alpha > 0$ and for all small enough $h \in \mathbb{R}^{dp}$,

$$d_\Pi(\phi(x)+h) - d_\Pi(\phi(x)) \geq \langle x^\star, h\rangle - \alpha|h|^2.$$

Let us fix an index $i \in \{1,\dots,p\}$. It follows that for all small enough $h_i \in \mathbb{R}^d$

$$\sqrt{d_\Pi(\phi(x))^2 + d_{C_i}(x+h_i)^2 - d_{C_i}(x)^2} - \sqrt{d_\Pi(\phi(x))^2} \geq \langle x_i^\star, h_i\rangle - \alpha|h_i|^2.$$

By a first order expansion, we get

$$\frac{d_{C_i}(x+h_i)^2 - d_{C_i}(x)^2}{2d_\Pi(\phi(x))} \geq \langle x_i^\star, h_i\rangle - \alpha'|h_i|^2,$$

with another numerical constant α'. Then, we obtain with another constant α'' and for all small enough $h_i \in \mathbb{R}^d$

$$\frac{d_{C_i}(x)}{d_\Pi}(\phi(x))\,(d_{C_i}(x+h_i) - d_{C_i}(x)) \geq \langle x_i^\star, h_i\rangle - \alpha''|h_i|^2.$$

If $x \in C_i$ then $d_{C_i}(x) = 0$ and so we deduce that $x_i^\star = 0$. If $x \notin C_i$ then by definition of the proximal normal cone,

$$\frac{d_\Pi(\phi(x))}{d_{C_i}(x)} x_i^\star \in \partial^P d_{C_i}(x) \subset S(0,1),$$

see Remark 4.8.
So $|x_i^\star| = d_{C_i}(x) d_\Pi(\phi(x))^{-1}$ and so $|x^\star| = 1$, which concludes the proof of (18).
Let us now finish the proof of the proposition. Thus

$$\partial^C f(x) \subset \sum_{i,\ x\notin C_i} \mathcal{E}_i(x) \subset \sum_{i,\ x\notin C_i} \frac{d_{C_i}(x)}{d_\Pi(\phi(x))}\partial^P d_{C_i}(x).$$

We set $z = (z_1,...,z_p) \in \mathbb{R}^{dp}$ with for all i, $z_i = \mathrm{P}_{C_i}(x)$ or equivalently $z = \mathrm{P}_\Pi(\phi(x))$. By Theorem 1.105 in [10],

$$\partial^P d_{C_i}(x) \subset \partial^P d_{C_i}(z_i) \cap S(0,1) = \left\{\frac{x-z_i}{|x-z_i|}\right\}.$$

Consequently, we have

$$\partial^C f(x) \subset \sum_{i,\ x\notin C_i} \mathcal{E}_i(x) \subset \left\{\sum_{i,\ x\notin C_i} \frac{d_{C_i}(x)}{d_\Pi(\phi(x))}\frac{x-z_i}{|x-z_i|}\right\} = \left\{\sum_{i,\ x\notin C_i} \frac{x-z_i}{|\phi(x)-z|}\right\}.$$

We finish the proof by choosing $y := \phi(x) - z \in \mathbb{R}^{dp}$. □

Remark 4.8. *Let $S \subset \mathbb{R}^d$ be a closed convex set and $x \notin S$, then $\partial^P d_S(x) = \partial^C d_S(x) \subset S(0,1)$. Indeed with $w := \mathrm{P}_S(x)$ and vectors $h = \epsilon(w-x)$ for small enough ϵ, we remark that $d_S(x+\epsilon(w-x)) = d_S(x) - \epsilon|w-x|$. Hence, by Definition 4.1, we obtain for every $x^\star \in \partial^P d_S(x)$*

$$-|h| \geq \langle x^\star, h\rangle - \alpha|h|^2.$$

By dividing by $|h|$ and letting ϵ go to 0, we deduce that $|x^\star| \geq 1$. We also conclude to $|x^\star| = 1$ since d_S is 1-Lipschitz.

Theorem 4.9. *There exist r and Θ such that for all $t \in [0,T]$, $\mathrm{q}_0 \in U(t)$ satisfying $d_{\tilde{Q}(t,\mathrm{q}_0)}(\mathrm{q}_0) \leq r/4$ and all $\mathrm{q} \in B(\mathrm{q}_0, r/4) \bigcap Q(t)$,*

$$d_{\tilde{Q}(t,\mathrm{q}_0)}(\mathrm{q}) \leq \Theta \sum_{i=1}^{p} d_{\tilde{Q}_i(t,\mathrm{q}_0)}(\mathrm{q}).$$

Indeed we can choose $\Theta = \frac{2\gamma\beta}{\alpha}$ and $r = \min(\frac{\rho}{6\beta}, \frac{\alpha}{4M\gamma})$.

Proof. Consider $r = \min(\frac{\rho}{6\beta}, \frac{\alpha}{4M\gamma})$. Let us fix $t \in [0,T]$, $\mathrm{q}_0 \in U(t)$ satisfying $d_{\tilde{Q}(t,\mathrm{q}_0)}(\mathrm{q}_0) \leq r/4$. Consequently there exists $\mathrm{q}_1 \in B(\mathrm{q}_0, r/4)$ such that $\mathrm{q}_1 \in \tilde{Q}(t,\mathrm{q}_0)$. We define a Lipschitz map $f := d_{F(\cdot)}(0)$ where F is given by (17) with $C_i = \tilde{Q}_i(t,\mathrm{q}_0)$. First we check the assumptions of Lemma 4.6 for the function f with $x_0 = \mathrm{q}_1$. Indeed $f(\mathrm{q}_1) = 0$ because $\mathrm{q}_1 \in \tilde{Q}(t,\mathrm{q}_0)$.
Let us consider $\mathrm{q} \in B(\mathrm{q}_1, r) \cap \tilde{Q}(t,\mathrm{q}_0)^c \cap Q(t)$, so $\mathrm{q} \in B(\mathrm{q}_0, 2r)$. By Proposition 4.7, $\partial^C f(\mathrm{q}) = \{\mathrm{q}^\star\}$ where

$$\mathrm{q}^\star := \sum_{i,\ \mathrm{q} \notin \tilde{Q}_i(t,\mathrm{q}_0)} \mathrm{p}_i^\star$$

with $\mathrm{p}^\star = \mathrm{p}/|\mathrm{p}|$ and $\mathrm{p} = \mathrm{P}_{F(\mathrm{q})}(0)$. Moreover for i satisfying $\mathrm{q} \notin \tilde{Q}_i(t,\mathrm{q}_0)$, $-\mathrm{p}_i^\star \in \mathrm{N}(C_i, \mathrm{q} + \mathrm{p}_i)$.
Let us define

$$J(t,\mathrm{q}) := \{j,\ g_j(t,\mathrm{q}_0) + \langle \nabla g_j(t,\mathrm{q}_0), \mathrm{q} - \mathrm{q}_0 \rangle < 0\} = \{j,\ \mathrm{q} \notin \tilde{Q}_j(t,\mathrm{q}_0)\}.$$

It is well-known that there also exist nonnegative reals $(\lambda_i)_{i \in J(t,\mathrm{q})}$ satisfying

$$\mathrm{q}^\star = \sum_{i \in J(t,\mathrm{q})} \lambda_i \nabla g_i(t,\mathrm{q}_0).$$

Hence by Assumption (A4)

$$\begin{aligned} |\mathrm{q}^\star| &= \left| \sum_{i \in J(t,\mathrm{q})} \lambda_i \nabla g_i(t,\mathrm{q}_0) \right| \\ &\geq \left| \sum_{i \in J(t,\mathrm{q})} \lambda_i \nabla g_i(t,\mathrm{q}) \right| - 2Mr \sum_{i \in J(t,\mathrm{q})} \lambda_i \end{aligned}$$

Since $\mathrm{q} + \mathrm{p}_i \in \mathrm{P}_{C_i}(\mathrm{q})$ and $\mathrm{q} \in Q(t)$, Proposition 3.6 yields $|\mathrm{p}_i| = d_{C_i}(\mathrm{q}) \leq \frac{2M}{\alpha} r^2$. Moreover for all $i \in J(t,\mathrm{q})$, $\mathrm{q} + \mathrm{p}_i \in \partial\tilde{Q}(t,\mathrm{q}_0)$ so we have by Assumption (A1)

$$\begin{aligned} g_i(t,\mathrm{q}_0) &= -\langle \nabla g_i(t,\mathrm{q}_0), \mathrm{q} + \mathrm{p}_i - \mathrm{q}_0 \rangle \leq \beta(|\mathrm{q} - \mathrm{q}_0| + |\mathrm{p}_i|) \\ &\leq 2\beta r \left(1 + \frac{Mr}{\alpha}\right). \end{aligned}$$

Hence by Assumption (A1),

$$g_i(t,\mathrm{q}) \leq 2\beta r \left(2 + \frac{Mr}{\alpha}\right).$$

Due to the choice of r, we deduce that $2\beta r \left(2 + \frac{Mr}{\alpha}\right) \leq \rho$ and thus $J(t,\mathrm{q}) \subset I_\rho(t,\mathrm{q})$. From

Assumptions (A1), (A4) and (A5'), we deduce that

$$
\begin{aligned}
|\mathrm{q}^\star| &\geq \gamma^{-1} \sum_{i\in J(t,\mathrm{q})} \lambda_i \, |\nabla g_i(t,\mathrm{q})| - 2Mr \sum_{i\in J(t,\mathrm{q})} \lambda_i \\
&\geq (\alpha\gamma^{-1} - 2Mr) \sum_{i\in J(t,\mathrm{q})} \lambda_i \\
&\geq \frac{\alpha}{2\gamma} \sum_{i\in J(t,\mathrm{q})} \lambda_i \\
&\geq \frac{\alpha}{2\gamma\beta} \sum_{i\in J(t,\mathrm{q})} \lambda_i \, |\nabla g_i(t,\mathrm{q}_0)| \\
&\geq \frac{\alpha}{2\gamma\beta} \sum_{i\in J(t,\mathrm{q})} |\mathrm{p}_i^\star| \\
&\geq \frac{\alpha}{2\gamma\beta} |\mathrm{p}^\star| = \frac{\alpha}{2\gamma\beta}.
\end{aligned}
$$

We can also apply Lemma 4.6 and we obtain that for all $\mathrm{q} \in B(\mathrm{q}_1, r/2) \supset B(\mathrm{q}_0, r/4)$

$$
d_{\tilde{Q}(t,\mathrm{q}_0)}(\mathrm{q}) \leq \Theta \left(\sum_{i\in J(t,\mathrm{q})} d_{\tilde{Q}_i(t,\mathrm{q}_0)}(\mathrm{q})^2 \right)^{1/2} \leq \Theta \sum_{i=1}^{p} d_{\tilde{Q}_i(t,\mathrm{q}_0)}(\mathrm{q}).
$$

□

References

[1] F.H. Clarke, R.J. Stern, and P.R. Wolenski. Proximal smoothness and the lower-C^2 property. *J. Convex Anal.*, 2:117–144, 1995.

[2] J.F. Edmond and L. Thibault. Relaxation of an optimal control problem involving a perturbed sweeping process. *Math. Program, Ser. B*, 104(2-3):347–373, 2005.

[3] J.F. Edmond and L. Thibault. BV solutions of nonconvex sweeping process differential inclusion with perturbation. *J. Differential Equations*, 226(1):135–179, 2006.

[4] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[5] A.D. Ioffe and J. V. Outrata. On metric and calmness qualification conditions in subdifferential calculus. *Set-valued Anal.*, 16:199–227, 2008.

[6] B. Maury. A time-stepping scheme for inelastic collisions. *Numerische Mathematik*, 102(4):649–679, 2006.

[7] B. Maury and J. Venel. Un modèle de mouvement de foule. In *ESAIM: Proc.*, volume 18, pages 143–152, 2007.

[8] B. Maury and J. Venel. A microscopic model of crowd motion. *C. R. Acad. Sci. Paris Ser. I*, 346:1245–1250, 2008.

[9] M.D.P. Monteiro Marques. Differential inclusions in nonsmooth mechanical problems, shocks and dry friction. Birkhäuser, Basel, 1993.

[10] B. S. Mordukhovich. Variational Analysis and Generalized Differentiation I: Basic Theory. Springer 2006.

[11] R.T. Rockafellar. Directionally Lipschitzian functions and subdifferential calculus. *Proc. London Math. Soc.*, 39:331–355, 1979.

[12] J. Venel. A numerical scheme for a whole class of sweeping process. To appear in *Numer. Math.*, available at http://arxiv.org/pdf/0904.2694, 2009.