Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Necessary Conditions of Optimality for a Time-Optimal Bi-level Sweeping Control Problem

Nathalie T. Khalil and Fernando Lobo Pereira *

* *SYSTEC, Faculty of Engineering, Porto University*
Institute for Systems and Robotics, Porto, Portugal
Email: khalil.t.nathalie@gmail.com, flp@fe.up.pt

Abstract: This article concerns a simple instance of an optimal control problem paradigm combining bi-level optimization with sweeping processes, initially investigated in Khalil and Lobo Pereira (2019). This class of problems arises, for instance, in structured crowd motion control problems in a confined space. We propose a specific class of time-optimal bi-level problem with sweeping process dynamics represented in terms of a *truncated normal cone* at the lower level. We establish the necessary optimality conditions of the Maximum Principle of Pontryagin type in the Gamkrelidze's form. Two techniques are at the core of the analysis: a) the *smooth approximation* of the low level sweeping control system, thereby avoiding the absence of Lipschitzianity inherent to the sweeping process, and, b) the *flattening* of the bi-level structure to a single level problem by using a exact penalization technique involving the value function of the low level problem to incorporate its inherent constraint in the bi-level structure. Necessary optimality conditions are applied to the resulting approximate flattened problem, and the main result of this article is obtained by passing to the limit.

Keywords: Optimal control theory, Sweeping process, Bi-level optimization, Gamkrelidze's necessary conditions of optimality.

1. INTRODUCTION

The aim of this article is to investigate a bi-level optimal problem coupled with an optimal sweeping control process. The formulation of an instance of this problem for the minimum time control of a structured crowd was given in Khalil and Lobo Pereira (2019), where existence and well-posedness results were also addressed. Now, we are concerned with deriving necessary conditions of optimality in the Gamkrelidze form of the Maximum Principle of Pontryagin, Gamkrelidze (1960), for a simple instance of the problem considered in the above reference.

More precisely, we consider the following high level problem

$$\begin{aligned}
(P_H(x_0,u))\quad & \text{Minimize } J_H(T,y;x_0,u)\\
\text{subject to } & \dot{y}(t)=v(t) \quad \text{a.e. in } [0,T]\\
& y(0)=y_0\in Q\subset\mathbb{R}^n, \text{ so that } y_0+Q_1\subset Q\\
& y(T)\in\bar{E},\ \bar{E}=\partial[(E+r_1B_1(0))\cap Q],\ E\subset\partial Q\\
& v\in\mathcal{V}:=\{v\in L^2([0,T];\mathbb{R}^n):\ v(t)\in V\}\\
& Q_1+y(t)\subset Q \qquad \forall\, t\in[0,T] \qquad (1)\\
& \text{and } (T,y) \text{ s.t. } \exists \text{ a solution to } P_L(T,y),
\end{aligned}$$

where y_0 is given, $J_H(T,y;x_0,u):=T$ (time-optimal), $V\subset\mathbb{R}^n$ is compact, the set E is closed, Q and Q_1 are spheres in $\mathbb{R}^n$ of radius r and r_1, respectively, with $r\gg r_1$, Q centered at q_0, ∂Q denotes the boundary of Q, and $(P_L(T,y))$ represents the following parametric low level problem whose dynamics involve a sweeping process.

$$\begin{aligned}
(P_L(T,y))\quad & \text{Minimize } J_L(x_0,u;T,y)\\
\text{subject to } & \dot{x}(t)\in f(x(t),u(t))-N^M_{Q_1+y(t)}(x(t)) \text{ a.e.} \qquad (2)\\
& x(0)=x_0\in Q_1+y_0\\
& u(t)\in\mathcal{U}:=\{u\in L^\infty([0,T];\mathbb{R}^m):\ u(t)\in U\}\\
& x(t)\in Q_1+y(t) \quad \forall t\in[0,T], \qquad (3)
\end{aligned}$$

where, for a given pair (T,y) solving the high level problem $(P_H(x_0,u))$, $J_L(x_0,u;T,y):=\int_0^T|u(t)|^2dt$ (control effort), $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $U\subset\mathbb{R}^m$ is compact, $N^M_A(z):=N_A(z)\cap MB_1(0)$, being $N_A(z)$ the normal cone to the set A at point z in the sense of Clarke (1990), $M>0$ is a given constant, and $B_1(0)$ the closed unit ball with center at the origin. We note that the consideration of a truncated normal cone at the low level sweeping process instead of the usual normal cone is due to the need to preserve the meaningfulness of the bi-level structure. Note that, with the usual normal cone, the high level control could drive Q_1 to the exit set E while the low level system would not need to do any extra effort to remain feasible, and, as a consequence, the bi-level structure would collapse to two independent optimal control problems, one of which with sweeping dynamics.

Unless there is a specific reason for so, we shall, for the sake of improving the presentation, drop the parameters (x_0,u) and (T,y), referring to the high and the low level problems as (P_H) and (P_L), respectively. Moreover, the bi-level problem emerging from the articulation of (P_H), and (P_L) is denoted by (P_B).

A pair $(x(\cdot), u(\cdot))$ is a *feasible* (or admissible) control process of (P_L) if $u(\cdot)$ is feasible control of (P_L), and $x(\cdot)$ is an arc satisfying the differential equation, the initial condition, together with (3) where y is such that, for some $T > 0$, (T, y) is a feasible pair of $(P_H(x_0, u))$. An *optimal solution* to (P_L) is a feasible pair of (P_L) minimizing the value of the cost functional $J_L(x_0, u; T, y)$ over all admissible pairs of $(P_L(T, y))$. A feasible quadruple of the bi-level dynamic control problem (P_B) is the collection of a feasible pair (T, y), and an optimal pair (x_0, u) to $(P_L(T, y))$. The feasible quadruple $(T, y; x_0, u)$ is optimal to (P_B) if $(T, y; x_0, u)$ minimizes the value of $J_H(T, y; x_0, u)$ among all admissible strategies of (P_B).

Note that the constraints (1), and (3) can, respectively, be expressed by the inequality constraints $h_H(y(t)) \leq 0$, and $h_L(x(t), y(t)) \leq 0$, where

$$h_H(y) := \frac{1}{2}(|y-q_0|^2-(r-r_1)^2), \;\; h_L(x, y) := \frac{1}{2}(|x-y|^2-r_1^2)$$

being $|\cdot|$ the Euclidean norm, and q_0 the center of Q.

Classical application scenarios of problems with sweeping control processes arise in, for instance, structured crowd motion control problems, operation of teams of drones offering complementary services in a shared confined space, smart material systems, nanoferro-electric, elasticity systems, among others, that exhibit some structure formed by a set of groups with different properties and confined in time-variant subsets contained in a larger confined space.

The simplest instance of a crowd motion control problem in a confined space can be formulated as a bi-level sweeping process problem like (P_B). The high level problem (P_H) affects the motion of the controlled low level sweeping process, by moving its constraint set $Q_1 \subset Q$ towards the exit set $E \subset \partial Q$. While the high level problem minimizes the final time T at which Q_1 reaches E, the low level problem minimizes its control effort to remain within $Q_1 + y(t) \subset Q \;\forall\; t \in [0, T]$ where $y(t)$ is a feasible arc to (P_H) which together with T, is a parameter to (P_L). Clearly, minimizing (P_L) constitutes a constraint to (P_H) only if the velocity set of the low level dynamics is adequately bounded. Thus, the proper formulation of this problem requires the truncation of the normal cone to $Q_1 + y(t)$ at $x(t)$. In this article, we investigate the corresponding necessary conditions of optimality in the Gamkrelidze's form for (P_B).

The article is organized as follows. In section 2, we present the main theorem of the paper: necessary optimality conditions of Maximum Principle Pontryaguin type in the Gamkrelidze form for the time-optimal bi-level sweeping control problem (P_B). The proof of the main theorem, which essentially relies on smoothing the sweeping dynamical system and on a flattening technique, is the subject of section 3. Section 4 includes a conclusion and outlines future development avenues.

Notation. We shall denote by $\partial^P\varphi$, $\partial\varphi$, and $\partial^C\varphi$, respectively, the proximal, the limiting, and the Clarke subdifferentials of the function φ. If φ is locally Lipschitz, then $\partial^C\varphi = \text{co}\;\partial\varphi$, where "co A" denotes the convex hull of the set A. Details on these concepts, and pertinent tools of nonsmooth analysis can be found, for instances, in Clarke (1990); Vinter (2010).

2. MAIN THEOREM

First, we state the assumptions to be imposed on the data of the problem.

- H1 $f(\cdot,\cdot)$ is continuous, $f(\cdot, u)$ is Lipschitz continuous for all $u \in U$, and there exists a constant $M_1 > 0$ such that $|f(x, u)| \leq M_1$ for all $(x, u) \in \mathbb{R}^n \times U$.
- H2 $f(x, U)$ is a closed and convex set for each $x \in \mathbb{R}^n$.
- H3 The control constraint sets U, and V are compact, and U is convex.
- H4 There exists $\delta{>}0$ s. t. $\delta B_1(0) \subset f(x, U)$, $\forall\, x \in \mathbb{R}^n$.
- H5 The constant M specifying the truncation level of normal cone has to satisfy $\bar{M} > M > \bar{m}$ where, $\forall\, \zeta \in N_{Q_1+y(t)}(x(t))$, $\forall\, t \in [0, T]$ s.t. $x(t) \in \partial(Q_1 + y(t))$,

$$\bar{M} := \min_{|\zeta|=1}\left\{\max_{u\in U}\{\langle\zeta, f(x(t), u)\rangle\} - \min_{v\in V}\{\langle\zeta, v\rangle\}\right\},$$

$$\bar{m} := \max_{|\zeta|=1}\left\{\min_{u\in U}\{\langle\zeta, f(x(t), u)\rangle\} - \max_{v\in V}\{\langle\zeta, v\rangle\}\right\}.$$

Assumption H4, and the need to consider a truncated cone in the lower level sweeping dynamics is crucial in order to preserve the bi-level structure of the problem. Further information can be found in (Khalil and Lobo Pereira, 2019, Section III).

Before stating the main theorem of this paper, we shall observe that the low level dynamic control system (2) coupled with (3) and the control constraint is equivalent to a control problem with a mixed constraint (i.e., joint constraint on the control and state variables), as follows:

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) - \dfrac{M}{r_1} u_0(t)(x(t) - y(t)) \\ h_L(x(t), y(t)) \leq 0, \;\; u_0(t) h_L(x(t), y(t)) = 0 \\ x(0) = x_0 \in Q_1 + y_0, \;\; u \in \mathcal{U}, \text{ and } u_0(t) \in [0, 1], \end{cases} \tag{4}$$

being the additional control $u_0{:}[0, T] \to [0, 1]$ a measurable function.

This reformulation of (P_L) becomes clear by observing that $(x(t) - y(t))$ is the gradient of h_L w.r.t. x, and, thus a direction in the normal cone of the state constraint set at its boundary point $x(t)$. Remark that, by just introducing an additional control u_0, a much more convenient single-valued and smooth representation of $\dot{x} \in f(x, u) - N^M_{Q_1+y}(x)$ is obtained.

Unfortunately, the mixed equality constraint in (4) is not regular and necessary conditions of optimality for problems featuring both irregular mixed constraint, and state constraints are not available.

Indeed, in Arutyunov et al. (2010, 2011), a maximum principle in the Gamkrelidze's form is established for an optimal control problem in the presence of mixed constraints, and pure state constraints. However, in these articles, the mixed constraints are regular. Dmitruk (2009) considers irregular mixed constraint but only problems without state constraints.

Let $\varphi : L^2([0, T]; \mathbb{R}^+) \times L^2([0, T]; \mathbb{R}^n) \to \mathbb{R}$ be the value function of (P_L) on the functional parameters of (P_H)

$$\varphi(\omega, v) = \min\{J_L(x_0, u; T, y){:}(x_0, u) \text{ feasible for } (P_L)\}, \tag{5}$$

being $v \in \mathcal{V}$ the control associated with the arc y, and ω an additional "control" yielding the final time T, and that featuren in the admissible control processes of an equivalent optimal control problem on the fixed time interval $[0, T^*]$ with t as an additional state component. This new problem is derived by the standard time variable change $t{:}[0,T^*]{\to}\mathbb{R}^+$, with $t(0){=}0$, and $t(\tau){=}\int_0^\tau \omega(s)ds$.

Problem (P_B) - the articulation of (P_H), and (P_L) as defined in section 1) - can, thus, be equivalently written (we do not relabel variables) as

$$\begin{aligned}
(P_B)\ \text{Minimize}\ & J_H(t, y; x_0, u) := t(T^*) \\
\text{subject to}\ & \dot{y} = v\omega,\ \dot{z} = |u|^2\omega,\ \dot{t} = \omega \\
& \dot{x} = f(x, u) + \bar{f}(x, y)u_0(x - y))\omega \qquad (6) \\
& y(0) = y_0,\ \ y(T^*) \in \bar{E},\ \ t(0) = 0 \\
& x(0) \in Q_1 + y_0,\ \ z(0) = 0 \\
& v{\in}\mathcal{V},\ u{\in}\mathcal{U},\ u_0{\in}[0, 1],\ \omega{\in}L^2([0, T^*]; \mathbb{R}^+) \\
& h_H(y) \le 0,\ \ h_L(x, y) \le 0\ \ \text{on } [0, T^*] \\
& z(T^*) - \varphi(\omega, v) \le 0. \qquad (7)
\end{aligned}$$

where $\bar{f}(x, y) = -\frac{M}{r_1}$ if $|x - y| = r_1$, and 0, otherwise. Remark that (P_B) is stated in terms of the new time parametrization in order to make the functional character (i.e., in an infinite dimensional space), and the "boundary" (i.e., not point-wise in time) nature of the interdependence between (P_H), and (P_L) in the variables (T, y) clearer and easier to handle. In what follows, the references to (P_B) should be always interpreted in the light of the relations $T{=}t(T^*)$, and $y(t) = y_0{+}\int_0^t v(\tau)d\tau$.

Let $(T^*, y^*, x^*, u^*, z^*, u_0^*)$ be a minimizer to (P_B). We require the following definition (cf. Ye (1997)).

Definition 2.1. (Partial Calmness). (P_B) is called partially calm at a minimizer $(T^*, y^*, x^*, u^*, z^*, u_0^*)$ with modulus ν if $\exists\ \nu \ge 0$ such that, for any feasible point (T, y, x, u, z, u_0), the following inequality holds:

$$\begin{aligned}
J_H(T, y; x_0, u) - J_H(T^*, y^*; x_0^*, u^*)& \\
+ \nu(z(T) - \varphi(\omega, v)) \ge 0.&
\end{aligned}$$

We are now ready to state the main result, the necessary optimality conditions in the Gamkrelidze's form for the bi-level optimization problem in the sense expressed in terms of (P_B). First, let

$$\begin{aligned}
H_H(y, x, v, u, q_H, q_L, \nu_H, \nu_L, r) = \langle q_H{-}\nu_H(y - q_0), v\rangle& \\
+\langle q_L{-}\nu_L(x - y), f(x, u)\rangle{+}\nu_L\langle x{-}y, v\rangle& \\
+\sigma(y, x, q_L, \nu_L) - r|u|^2&
\end{aligned}$$

where y, x, v, u, q_H, q_L are in $\mathbb{R}^n$, ν_H, ν_L, r are nonnegative scalars, and

$$\sigma(y, x, q_L, \nu_L){=}\sup_{w\in -N^M_{Q_1}(x-y)} \{\langle q_L{-}\nu_L(x - y), w\rangle\}.$$

Alternatively, $\sigma(y, x, q_L, \nu_L){=}\max\{0, \dfrac{M}{r_1}(\nu_L r_1^2{-}\langle q_L, x{-}y\rangle)\}$ if $|x - y| = r_1$, and 0 if $|x - y| < r_1$.

Theorem 2.2. Let H1-H5 hold. Let (T^*, y^*, x_0^*, u^*) be a solution to (P_B), assumed to be partially calm at its solution with modulus ν. Then, there exists a set of multipliers $(q_H, q_L, \nu_H, \nu_L, \lambda, r)$ with q_H, and q_L in $AC([0, T^*]; \mathbb{R}^n)$, ν_H, and ν_L in $NBV([0, T^*]; \mathbb{R}^+)$ are non-increasing, being ν_H constant on $\{t{\in}[0, T^*]{:}\,|y - z|{>}r_1\,\forall z{\in}\partial Q\}$, and ν_L constant on $\{t{\in}[0, T^*]{:}\,|y - x|{<}r_1\}$, and λ, and r, with $r = \lambda\nu$ are non-negative numbers, satisfying the following conditions:

1. Nontriviality. $|(q_H, q_L, \nu_H, \nu_L)| + \lambda + r \ne 0$
2. Adjoint Equations

$$\begin{aligned}
-\dot{q}_H(t){\in}{-}\langle\nu_H(t){+}\ \nu_L(t), v^*(t)\rangle{+}\nu_L(t)f(x^*(t), u^*(t)) \qquad (8)& \\
+\partial_y\sigma(y^*(t), x^*(t), q_L(t), \nu_L(t))\ \ [0, T^*] - \text{a.e.}& \\
-\dot{q}_L(t){\in}\partial_x\langle q_L(t){-}\nu_L(t)(x^*(t){-}y^*(t)), f(x^*(t), u^*(t))\rangle \qquad (9)& \\
+\nu_L v^*(t){+}\partial_x\sigma(y^*(t), x^*(t), q_L(t), \nu_L(t))\ [0, T^*] - \text{a.e.}&
\end{aligned}$$

3. Boundary Conditions

$$\begin{aligned}
&q_H(0) \in \mathbb{R}^n, \text{ and } q_L(0) = \nu_L(0)(x^*(0) - y_0) \\
&q_H(T^*) \in -N_{\bar{E}}(y^*(T^*)) + \nu_H(T^*)(y^*(T^*) - q_0) \\
&\qquad\qquad -\nu_L(T^*)(x^*(T^*) - y^*(T^*)) \\
&q_L(T^*) = \nu_L(T^*)(x^*(T^*) - y^*(T^*))
\end{aligned}$$

4. Time Transversality Condition. For all $t \in [0, T^*]$,

$$H_H(y^*, x^*, v^*, u^*, q_H, q_L, \nu_H, \nu_L, r)(t) = \lambda \qquad (10)$$

5. Maximum Condition on the low level control $u^*(t)$ maximizes on U, $[0, T^*]$-a.e. the mapping

$$u{\to}\langle q_L(t){-}\nu_L(t)(x^*(t){-}y^*(t)), f(x^*(t), u)\rangle{-}r|u|^2 \qquad (11)$$

6. Maximum Condition on the high level control

$$(0, q_H{-}\nu_H(y^*{-}q_0)){\in}{-}r\partial\varphi(T^*, v^*){+}\{0\}{\times}N_{\mathcal{V}}(v^*) \qquad (12)$$

3. PROOF OUTLINE OF THEOREM 2.2

The proof of theorem 2.2 is organized in four steps.

The first one addresses the challenge inherent to the presence of a sweeping process in the dynamics of the low level problem which entails the fact that the dynamics fails to be Lipschitz continuous w.r.t. the state variable. In order to overcome this difficulty, we consider a sequence of auxiliary problems $\{(P_L^k)\}$ approximating the low level problem (P_L) such that the dynamics of each one of the (P_L^k)'s is Lipschitz continuous w.r.t. the state variable by using a construction involving an exponential function that depends on the parameter k. Our construction to the approximate low level dynamics featuring the truncated normal cone necessarily differs from the one in de Pinho et al. (2019); Zeidan et al. (2019) where the usual normal cone is considered. Moreover, it is important to note that, in our approximation to the low level problem, the state constraint condition cannot be discarded from the formulation, contrarily to the work in (de Pinho et al., 2019, Lemma 1) or (Zeidan et al., 2019, Lemma 4.2) where it is proved that the trajectory always remains in the state constraint set, enabling its elimination in the smooth approximated problem.

The second step consists in flattening the approximating the bi-level problem by using an exact penalization technique to obtain a standard "one-level" optimal control problem. This is done by replacing the low level problem in the bi-level structure by inserting in the high level optimal control problem an additional constraint involving the low level problem value function with the high level controls as arguments. Then, a partial calmness condition of the approximating problem plays a key role in constructing a related optimal control problem by using a penalization technique consisting in eliminating the additional constraint involving the value function, while adding to the cost functional an appropriate penalty term.

Although, the existence of solution to each (P_L^k) can be asserted whenever (T, y) is such that the set of feasible control processes is not empty, the fact is that we can not relate it with the given reference control process (u^*, x^*). This leads to the need of a third step, in which the existence of a solution to a suitably perturbed version of this approximated bi-level problem (P_B^k) through the application of Ekeland's variational principle is asserted. Moreover, it is also shown the existence of a sub-sequence of problems whose solutions converge to the reference optimal control process considered in the main result statement.

Given the properties of each one of the problems of the obtained sequence of approximate state constrained single level optimal control problems, we apply the maximum principle of Pontryaguin in the Gamkrelidze's form, cf. (Arutyunov et al., 2011, Theorem 3.1) or Arutyunov et al. (2010). Finally, we show that the necessary conditions of optimality to the original bi-level problem stated in our main result, Theorem 2.2, can be recovered by passing to the limit.

3.1 Approximation of the Lower Level Dynamics

Here, we construct a sequence of conventional control processes approximating the feasible sweeping control process in (P_L) by a smooth penalization being inspired by the technique recently introduced in de Pinho et al. (2019); Zeidan et al. (2019). However, the approximation approach has to be adapted to our context since we consider: 1) a truncated normal cone instead of the complete normal cone; and 2) a constant set moving in time $Q_1+y(t)$ instead of a given constant set.

The approximate low level dynamic system that we consider is as follows

$$(D^k)\begin{cases} \dot{x} = f^k(x, y, u, u_0) \;\; \mathcal{L}-\text{a.e. in } [0, T] \\ h_L(x, y) \leq 0 \quad \text{for all } t \in [0, T] \\ x(0) = x_0 \in Q_1 + y_0, \\ u \in \mathcal{U}, \;\; u_0 \in [0, 1] \;\; \mathcal{L}-\text{a.e. in } [0, T] \end{cases}$$

where

$$f^k(x, y, u, u_0) := f(x, u) - u_0 \min\left\{\frac{M}{r_1}, \gamma_k e^{\gamma_k h_L(x,y)}\right\}(x-y)$$

and $\{\gamma_k\}$ is a sequence such that

$$\lim_{k\to\infty} \gamma_k = \infty, \text{ and, for all } k, \; \gamma_k > \frac{M}{r_1}. \tag{13}$$

Let $\mathcal{F}_L^k(T, y)$ be the set of all processes (x, u) s.t. (T, y, x, u) is feasible for (D^k) and for (P_H). From assumptions H4, and H5, it is easy to conclude the existence of a value of M such that, for k sufficiently large, there are pairs (T, y) for which $\mathcal{F}_L^k(T, y) \neq \emptyset$. Moreover, from this, and the data of the low level problem, it is straightforward to conclude that the associate set of trajectories is compact. This entails that, for any such pair (T, y), there exists a solution (x, u) - the dependence on (T, y) is omitted - to the k-approximation low level problem

$$(P_L^k) \text{ Minimize } J_L(x_0, u; T, y)$$
$$\text{subject to conditions } (D^k).$$

It follows from the proposed approximation scheme, that

$$\lim_{k\to+\infty} u_0 \min\left\{\frac{M}{r_1}, \gamma_k e^{\gamma_k h_L(x,y)}\right\}(x-y) \in N_{Q_1}^M(x-y)$$

uniformly in x for any (T, y) s.t. $\mathcal{F}_L^k(T, y) \neq \emptyset$, and, thus, for such (T, y)'s, $f^k(x, y, u, u_0) \to f(x, u) - w(u_0)$ uniformly in (x, u) for some $w(u_0) \in N_{Q_1+y}^M(x)$.

Space limitations prevents the inclusion of the proofs of a number of auxiliary results required to prove the main theorem.

Proposition 1. Consider (P_B) and that $H1 - H5$ are in force. Then, $\exists \mathcal{N}^*$ a neighborhood of (T^*, y^*, x^*, u^*), such that, $\forall\; (T, y) \in \Pi_H(\mathcal{N}^*)$ (the high level components of $\mathcal{N}^*$), we have that $\mathcal{F}_L^k(T, y) \neq \emptyset$. Moreover, $\exists\; \{(T^k, y^k, x^k, u^k)\}$, s.t., for sufficiently large k, $(x^k, u^k) \in \mathcal{F}_L^k(T^k, y^k)$, $(x^k, u^k) \to (x^*, u^*)$ in $AC \times L^\infty$, and $(T^k, y^k) \to (T^*, y^*)$ in $\mathbb{R} \times AC$.

Denote the value function φ, given in (5), for (P_L^k) by φ^k.

Proposition 2. Given the data of (P_H), (P_L), and their articulation made precise in (P_B), it is always possible to choose a sequence $\{(T^k, y^k, x^k, u^k)\}$ satisfying the properties in Proposition 1 such that, for a given (T^k, y^k), the solution $(\bar{x}^k, \bar{u}^k)$ to (P_L) is in the interior of its solution set, and, thus, (w^k, v^k) is in the interior of the domain of φ^k, being the latter given by $\{(\omega, v) : \varphi^k(\omega, v) < \infty\}$.

Proposition 3. $\varphi^k(\omega, v)$ is locally Lipschitz continuous at any point in the interior of its domain.

Since we consider sequences approximating the solution in the interior of the domain of φ^k, we adapt the Theorem 2.3 of Ye (1997) to our Hamilton-Pontryagin function in the Gamkrelidze form, cf. Khalil and Lobo Pereira (2020), to compute its subgradients. Now, let $\bar{p} = (p_H, p_L)$, $\bar{\mu} = (\mu_H, \mu_L)$, $(\dot{y}, \dot{x}) = (v, f(x, u) - u_0 c(\gamma_k, y, x)(x-y))\omega$, with $c(\gamma_k, y, x) = \min\{\frac{M}{r_1}, \gamma_k e^{\gamma_k h_L(x,y)}\}$. Here, the Gamkrelidze term Γ in the Pontryagin-Hamilton function in Arutyunov et al. (2011) is $\Gamma = (\Gamma_H, \Gamma_L)$, where $\Gamma_L(\cdot) = \langle x-y, f(x, u) - u_0 c(\gamma_k, y, x)(x-y-v)\rangle\omega$ and $\Gamma_H(\cdot) = \langle y - q_0, v\rangle\omega$. Then, $H_L^k(\cdot) = \bar{H}_L^k(\cdot)\omega$ where

$$\bar{H}_L^k(y, x, v, \bar{p}, \bar{\nu}, \lambda) := \langle p_H - \mu_H(y - q_0), v\rangle + \mu_L\langle x-y, v\rangle \tag{14}$$
$$+\sigma_1(y, x, p_L, \mu_L, \lambda) + \sigma_2^k(y, x, q_L, \mu_L) - \lambda$$

where $\sigma_1(\cdot) = \sup_{u\in U}\{\langle q_L - \nu_L(x-y), f(x, u)\rangle - \lambda|u|^2\}$, and $\sigma_2^k(\cdot) = c(\gamma_k, y, x)\max\{0, \nu_L|x-y|^2 - \langle q_L, x-y\rangle\}$.

Proposition 4. Let $\Psi^k(\omega, v) = \{x : (x, u) \in \mathcal{F}_L^k(T, y)\}$. Then,

$$\partial\varphi^k(\omega, v) := \bigcup_{x\in\Psi^k(\omega,v)} \{\zeta \in L^2([0, T]: \mathbb{R}^{n+1}) : \exists\, \bar{p} \in AC \text{ s.t.}$$
$$(-\dot{\bar{p}}, -\zeta, \dot{y}, \dot{x}) \in \partial_{(y,x,\omega,v,\bar{p})} H_L^k(x, \omega, v, \bar{p}, \bar{\mu}, \lambda) \tag{15}$$
$$\text{and } (\bar{p}(0), -\bar{p}(T)) \in P_L\}$$

$P_L = \{(\bar{p}(0), \bar{p}(T)) : p_H(0) \in \mathbb{R}^n, -p_H(T) \in N_{\bar{E}}(y(T)), (\bar{p}_L(0), -\bar{p}_L(T)) \in N_{Q_1\times Q_1}(x(0) - y_0, x - y(T))\}$.

3.2 Flattening the Bi-Level Approximation Problem

We start by stating the flattened bi-level approximation problem cast as an equivalent optimal control problem on a fixed time interval $[0, T^*]$, by reparameterizing time in τ, while considering the previous time variable $t(\tau)$ as an additional state component. We do not relabel other variables.

$(\bar{P}^k_B)$ Minimize $t(T^*)$

subject to:

$$(\bar{D}^k_B)\begin{cases} \dot{x} = f^k(y, x, u, u_0)\omega, \;\; \dot{y} = v\omega, \;\; [0, T^*] - \text{a.e.} \\ x(0) = x_0 \in Q_1{+}y_0, \;\; y(0) = y_0, \;\; y(T^*) \in \bar{E}, \\ \dot{z} = |u|^2\omega, \;\; \dot{t} = \omega, \;\; [0, T^*] - \text{a.e.} \\ z(0) = 0, \;\; t(0) = 0, \\ u \in \mathcal{U}, \; v \in \mathcal{V}, \\ u_0{\in}L^1([0, T^*]; [0, 1]), \; \omega{\in}L^2([0, T^*]; \mathbb{R}^+), \\ h_H(y) \le 0, \quad h_L(x, y) \le 0 \quad \forall t \in [0, T^*], \\ z(T^*) - \varphi^k(\omega, v) \le 0 \end{cases}$$

The solution to this problem is not known, and, in order to pursue, we use Ekeland's variational principle, (cf. (Vinter, 2010, Theorem 3.3.1)). From Proposition 1, it follows $\exists \, \{\chi^k\}$, where $\chi^k = (y, x, z, t, v, u, u_0, \omega)^k$ feasible for $(\bar{D}^k_B)$ such that $\chi^k \to \chi^*$ where χ^* corresponds to $(y^*, x^*, v^*, u^*, u^*_0, T^*)$ solution to (P_B). This entails that, $\exists \, \bar{k}$ sufficiently large s.t., for $k > \bar{k}$, $\varepsilon_k = |t^k(T^*) - T^*| \to 0$ as $k \to \infty$.

Fix some $k \in \mathbb{N}$ sufficiently large. Let $\vartheta = (x_0, \varpi) \in \mathbb{R}^n{\times}L^2([0, T^*]; \mathbb{R}^n){\times}L^1([0, T^*]; \mathbb{R}^{n+1}){\times}L^2([0, T^*]; \mathbb{R}^+)$ with $\varpi = (v, u, u_0, \omega)$ and $\Upsilon = \{\vartheta : \vartheta \text{ is feasible for } (\bar{D}^k_B)\}$.

It is not difficult to see that (Υ, Δ), with $\Delta{:}\Upsilon{\times}\Upsilon{\to}[0, \infty)$ defined by $\Delta(\vartheta^a, \vartheta^b) := |x^a_0 - x^b_0| + \int_0^{T^*} |(v, u, u_0, 1)^a w^a - (v, u, u_0, 1)^b w^b| d\tau$, is a complete metric space. Notice also that $J_H(t(T^*), y; x_0, u) := t(T^*)$ is lower semi-continuous. Moreover, from the above, χ^k is a ε_k-minimizer to $(\bar{P}^k_B)$. Thus Ekeland's Variational Principle yields the existence of $\bar{\chi}^k$ associated with $(\bar{x}^k_0, \bar{v}^k)$ s.t. $\Delta((\bar{x}^k_0, \bar{\varpi}^k), (x^k_0, \varpi^k)) < \sqrt{\varepsilon_k}$ and solves the auxiliary problem:

$$(\widetilde{P}^k_B) \text{ Minimize } t(T^*){+}\sqrt{\varepsilon_k}\Delta((\bar{x}^k_0, \bar{\varpi}^k), (x_0, \varpi)) \quad \text{subject to } (\bar{D}^k_B)$$

3.3 Partial calmness of the approximation bi-level problem

To proceed with the proof, we still need to deal with the degeneracy inherent to the last functional inequality constraint in $(\bar{D}^k_B)$. In fact, the standard constraint qualifications, such as Mangasarian-Fromovitz or linear independence constraint, are too strong to hold. However, if the problem satisfies the *partial calmness condition*, then an approach based on a exact penalization, whereby the original problem can be replaced by an equivalent one (i.e., with same solution) without his last inequality constraint, but whose high-level objective function features a penalty term forcing its satisfaction. This is a key idea used in Ye and Zhu (1995); Ye (1997) to obtain necessary conditions of optimality, and, subsequently used in Dempe et al. (2014, 2007); Benita and Mehlitz (2016); Benita et al. (2016) in a wide variety instances, encompassing from finite to infinite dimensional low level problems. Here, the coupling parameter is control pair (w, v).

Proposition 5. The partial calmness of (P_B) entails the one for $(\bar{P}^k_B)$. Thus, $\forall \varepsilon_k{\downarrow}0$, $(\widetilde{P}^k_B)$ is also partially calm.

Theorem 3.1. Let $\bar{\chi}^k$ solve $(\widetilde{P}^k_B)$. Assume that

$$(\bar{\omega}^k, \bar{v}^k) \text{ is in interior of the domain of } \varphi^k. \tag{16}$$

Then, $(\widetilde{P}^k_B)$ is partially calm if and only if $\bar{\chi}^k$ solves

$$(\mathbf{P}^k_B) \text{ Minimize } \; t(T^*) + \sqrt{\varepsilon_k}\Delta((x_0, v), (\bar{x}^k_0, \bar{v}^k)) + \nu^k(z(T^*) - \varphi^k(\omega, v))$$

subject to $(\bar{D}^k_B)$ without the last inequality.

3.4 Necessary conditions of optimality for $(\mathbf{P}^k_B)$.

Clearly, the cost functional of $(\mathbf{P}^k_B)$ makes it a nonstandard optimal control problem. Thus, a path to the derivation of the necessary conditions of optimality involves casting the problem as a nonlinear programming problem in appropriate infinite dimensional spaces, apply the Fermat principle to the associated Lagrangian (the required assumptions hold), and, then, decode the obtained conditions in terms of the data of $(\mathbf{P}^k_B)$. This approach was adopted, for example, in the proof of Theorem 3.2 in Ye (1997). Albeit rather technical, these steps are straightforward, and, thus, also due to lack of space, we omit them, and jump directly to the resulting conditions that we choose to involve the Gamkrelidze form of the Pontryagin-Hamilton function. Given the complexity of the expressions that follow, consider $c(\gamma_k, y, x)$ as in subsection 3.1, and denote any function $l(a, b)$ by either $\bar{l}^k$ or by $\bar{l}^k(a)$ if its arguments are, respectively, either $(\bar{a}^k, \bar{b}^k)$ or $(a, \bar{b}^k)$. Recall $f^k(y, x, u, u_0) = f(x, u) - u_0 c(\gamma_k, y, x)(x - y)$ and let $r^k = \lambda^k \nu^k$, and

$$\widetilde{H}^k(y, x, z, v, u, u_0, q^k_H, q^k_L, \nu^k_H, \nu^k_L, \lambda^k, r^k) = -r^k|u|^2 - \lambda^k + \langle q^k_H{-}\nu^k_H(y - q_0), v\rangle{+}\langle \nu^k_L(x{-}y), v\rangle + \langle q^k_L{-}\nu^k_L(x{-}y), f^k(y, x, u, u_0)\rangle$$

Then, the necessary conditions of optimality for $(\mathbf{P}^k_B)$ can be written as follows.

If $(\bar{y}^k, \bar{x}^k, \bar{z}^k, \bar{t}^k, \bar{u}^k, \bar{u}^k_0, \bar{\omega}^k)$ solves $(\mathbf{P}^k_B)$, then, there exists a multiplier $(q^k_H, q^k_L, \nu^k_H, \nu^k_L, \lambda^k, r^k) \in AC([0, T^*]; \mathbb{R}^{2n}) \times C([0, T^*]; \mathbb{R}^{2n}) {\times} (\mathbb{R}^+)^2$ satisfying

1. Nontriviality. $|(q^k_H, q^k_L, \nu^k_H, \nu^k_L)| + \lambda^k + r^k \neq 0$
2. Adjoint equations (holding $[0, T^*]$-a.e.)
$$-\dot{q}^k_H {\in} (-\langle \nu^k_H, \bar{v}^k\rangle {+} \partial_y \langle q^k_L {-} \nu^k_L(\bar{x}^k {-} \bar{y}^k), \bar{f}^k\rangle {-} \nu^k_L \bar{v}^k)\bar{\omega}^k$$
$$-\dot{q}^k_L \in (\partial_x \langle q_L - \nu_L(\bar{x}^k - \bar{y}^k), \bar{f}^k\rangle {+} \nu^k_L \bar{v}^k)\omega^k$$
3. Boundary Conditions
$$q^k_H(0){\in}\mathbb{R}^n, \text{ and } q^k_L(0){\in}N_{Q_1+y_0}(\bar{x}^k(0)){-}\lambda^k\sqrt{\varepsilon_k}\xi^k_x$$
$$q^k_H(T^*) \in -N_{\bar{E}}(\bar{y}^k(T^*)) + \nu^k_H(T^*)(\bar{y}^k(T^*)) - \nu^k_L(T^*)(\bar{x}^k(T^*) - \bar{y}^k(T^*))$$
$$q^k_L(T^*) = \nu^k_L(T^*)(\bar{x}^k(T^*) - \bar{y}^k(T^*))$$
4. Maximum Conditions. The additive structure of $(\mathbf{P}^k_B)$ and the fact that $\omega > 0$ allows to separate the

maximum conditions for each one of the controls. Thus, we may write

- $\bar{u}^k$ maximizes on U, $[0, T^*]$-a.e. the mapping

$$u \to \langle q_L^k - \nu_L^k(\bar{x}^k - \bar{y}^k), f(\bar{x}^k, u)\rangle - r^k|u|^2 - \lambda^k\sqrt{\varepsilon_k}|u - \bar{u}^k| \quad (17)$$

- Maximum Condition on v.

$$0 \in (-(q_H^k - \nu_H^k(\bar{y}^k - q_0)) - \nu_L^k(\bar{x}^k - \bar{y}^k) - \lambda^k\sqrt{\varepsilon_k}\xi_v^k)\bar{\omega}^k - r^k\partial_v\varphi(\bar{\omega}^k, \bar{v}^k) + N_{\mathcal{V}}(\bar{v}^k) \quad (18)$$

- Maximum Condition on ω.

$$0 \in -\bar{\tilde{H}}^k + \lambda^k\sqrt{\varepsilon_k}(|\bar{u}^k| + |\bar{v}^k| + |\bar{u}_0^k| + 1)\xi_\omega^k - r^k\partial_\omega\varphi(\bar{\omega}^k, \bar{v}^k) \quad (19)$$

- $\bar{u}_0^k$ maximizes on $[0, 1]$, $[0, T^*]$-a.e. the mapping

$$u_0 \to -c(\gamma_k, \bar{y}^k, \bar{x}^k)\langle q_L^k - \nu_L^k(\bar{x}^k - \bar{y}^k), \bar{x}^k - \bar{y}^k\rangle u_0 - \lambda^k\sqrt{\varepsilon_k}|u_0 - \bar{u}_0^k| \quad (20)$$

In the above $\xi_a^k \in \partial|a - \bar{a}^k|_{a=\bar{a}^k}$. Note that $|\xi_a^k| = 1$.

3.5 Limit Taking when $k \to +\infty$.

In this last step we outline how our main result, Theorem 2.2 follows from the necessary conditions of optimality for $(\mathbf{P}_B^k)$, by taking the appropriate limit as $k \to \infty$. First note that it is not difficult to conclude that, by construction, there exists a subsequence $(\bar{y}^k, \bar{x}^k, \bar{u}^k, \bar{u}_0^k, \bar{\omega}^k)$ converging in an appropriate sense to $(y^*, x^*, u^*, \omega^*)$, with $\int_0^{T^*}\omega^*(\tau)d\tau = T^*$, the solution to the original bi-level problem (P_B). Note also that the sequence of functions $\{(q_H^k, q_L^k)\}$ is uniformly bounded, and equi-continuous, and $\{(\dot{q}_H^k, \dot{q}_L^k)\}$ is uniformly bounded in L^1. Thus, by standard analysis, and measure theory results, (see, e.g., de Pinho et al. (2019)), there are subsequences of $\{(q_H^k, q_L^k)\}$, converging uniformly to (q_H, q_L). Standard arguments entail the existence of a subsequence of $\{(\nu_H^k, \nu_L^k)\}$ converging pointwisely to $\{(\nu_H, \nu_L)\}$. By scaling uniformly along the sequence we ensure not only the non-triviality condition $\forall\ k$, but also that $\{(\lambda^k, r^k)\}$ take values in compact sets, and, thus, some subsequence converges to (λ, r). The coupling of these sequences via the above optimality conditions enable the extraction of a joint subsequence converging for the desired limits for all the above sequences. From now on, we always consider as a reference this subsequence. From the above, we conclude that (q_H, q_L) satisfy the respective adjoint equations and boundary conditions of our main result. Note that, by taking the limit as $k \to \infty$ of (20), we obtain the function σ defined in 2. Remark also, that the same limit taking in (17) leads to (11). By taking the limits in (18), and (19) along the reference subsequence, we obtain, respectively, $q_H - \nu_H(y^* - q_0) \in -r\partial_v\varphi(\omega^*, v^*) + N_{\mathcal{V}}(v^*)$, and $H_H(t) - \lambda \in r\partial_\omega\varphi(\omega^*, v^*)$. A standard contradiction argument leads to the transversality condition (10) and, at the same time to (12).

4. CONCLUSION

We investigated a specific bi-level optimal control problem whose low level dynamics are given by a sweeping process. Necessary conditions of optimality are proved under mild assumptions by a constructive method involving the approximation of the sweeping process by a conventional one. Future work encompasses generalizations of this problem to multi-objective functions at the low level. This context requires the investigation of more sophisticated solution concepts what, in itself, raises interesting challenges.

ACKNOWLEDGEMENTS

The support of FCT R&D Unit SYSTEC, ref. POCI-01-0145-FEDER-006933, ERDF COMPETE2020 FCT PT2020, STRIDE, ref. NORTE-01-0145-FEDER-000033, ERDF NORTE 2020, and MAGIC, ref. POCI-01-0145-FEDER-032485, FEDER COMPETE20 POCI and PIDDAC via FCT/MCTES is acknowledged.

REFERENCES

Arutyunov, A.V., Karamzin, D.Y., and Lobo Pereira, F. (2011). The maximum principle for optimal control problems with state constraints by RV Gamkrelidze: revisited. *Journal of Optimization Theory and Applications*, 149(3), 474–493.

Arutyunov, A., Karamzin, D.Y., and Lobo Pereira, F. (2010). Maximum principle in problems with mixed constraints under weak assumptions of regularity. *Optimization*, 59(7), 1067–1083.

Benita, F., Dempe, S., and Mehlitz, P. (2016). Bilevel optimal control problems with pure state constraints and finite-dimensional lower level. *SIAM Journal on Optimization*, 26(1), 564–588.

Benita, F. and Mehlitz, P. (2016). Bilevel optimal control with final-state-dependent finite-dimensional lower level. *SIAM Journal on Optimization*, 26(1), 718–752.

Clarke, F.H. (1990). *Optimization and nonsmooth analysis*, volume 5. SIAM.

de Pinho, M.d.R., Ferreira, M., and Smirnov, G. (2019). Optimal control involving sweeping processes. *Set-Valued and Variational Analysis*, 27(2), 523–548.

Dempe, S., Dutta, J., and Mordukhovich, B. (2007). New necessary optimality conditions in optimistic bilevel programming. *Optimization*, 56(5-6), 577–604.

Dempe, S., Mordukhovich, B.S., and Zemkoho, A.B. (2014). Necessary optimality conditions in pessimistic bilevel programming. *Optimization*, 63(4), 505–533.

Dmitruk, A. (2009). On the development of Pontryagin's maximum principle in the works of A.Ya. Dubovitskii and A.A. Milyutin. *Control & Cybernetics*, 38, 923–957.

Gamkrelidze, R.V. (1960). Optimal control processes for bounded phase coordinates. *Izvestiya Rossiiskoi Akad. Nauk. Ser. Matematicheskaya*, 24(3), 315–356.

Khalil, N. and Lobo Pereira, F. (2020). Necessary conditions of optimality in the Gamkrelidze's form for state constrained problems with differential inclusion. *IEEE Control Systems Letters*, accepted.

Khalil, N.T. and Lobo Pereira, F. (2019). A framework for the control of bilevel sweeping processes. In *IEEE 58th Conference on Decision and Control*, 6175–6180.

Vinter, R. (2010). *Optimal control*. Springer S & B Media.

Ye, J.J. (1997). Optimal strategies for bilevel dynamic problems. *SIAM Journal on Control and Optimization*, 35(2), 512–531.

Ye, J. and Zhu, D. (1995). Optimality conditions for bilevel programming problems. *Optimization*, 33(1), 9–27.

Zeidan, V., Nour, C., and Saoud, H. (2019). A nonsmooth maximum principle for a controlled nonconvex sweeping process. *preprint*.